

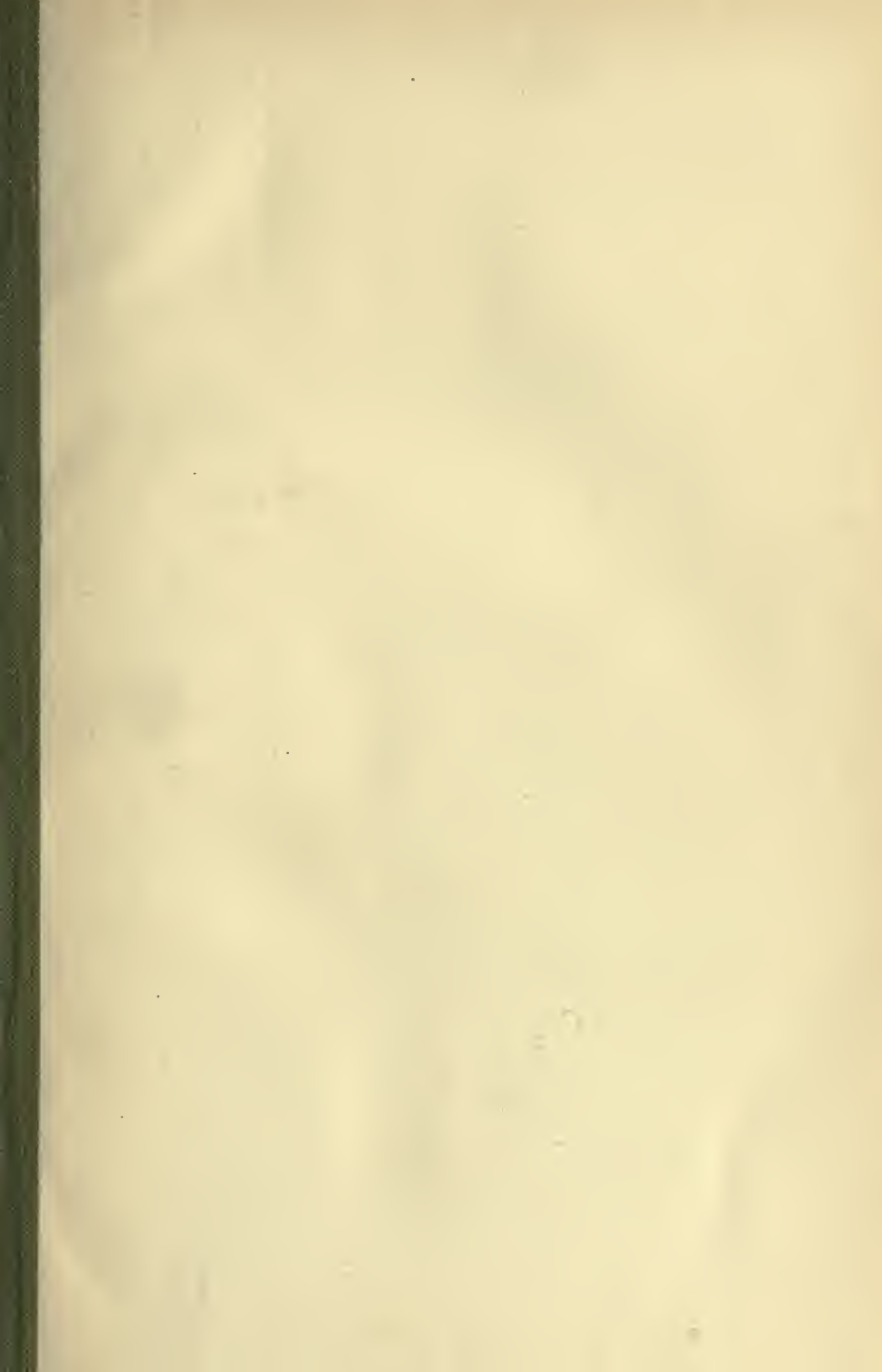
UNIVERSITY OF TORONTO



3 1761 00289322 0

UNIV OF  
TORONTO  
LIBRARY







Digitized by the Internet Archive  
in 2007 with funding from  
Microsoft Corporation







Philos  
A

I

# ARCHIVES OF PSYCHOLOGY

EDITED BY  
R. S. WOODWORTH

VOLUME VII  
1922

(a-e)  
226736  
7. 11. 28

NEW YORK  
1922



BF

21

A7

v.7

## CONTENTS.

Serial Number		Pages
47	The Effects of Practice on Judgments of Absolute Pitch. Evelyn Gough.	94
48	An Experimental Study of Silent Thinking. Ruth S. Clark.	101
49	Some Empirical Tests in Vocational Selection. Herbert W. Rogers.	46
50	Adenoids and Diseased Tonsils. Margaret Cobb Rogers.	70
51	An Experimental Study of the Factors and Types of Voluntary Choice. Alfred H. Martin.	115
52	Some Well Known Mental Tests Evaluated and Compared. Dorothy R. Morgenthau.	54



U 1

## TABLE OF CONTENTS

CHAPTER I. Introduction.	
1. Statement of the problem .....	7
2. Review of experimental studies and theories in regard to memory for tones .....	7
CHAPTER II. Observers .....	26
CHAPTER III. Experiments .....	27
1. Test I .....	27
a. Practice in the identification of the 88 piano notes (Group 1).	
b. Identification time.	
2. Test II .....	28
Practice in the identification of piano notes beginning with a single octave and increasing the range by octaves.	
3. Test III .....	28
Practice in the identification of violin notes.	
4. Test IV .....	29
Identification of pipe-organ notes: flute, diapason, oboe and viola stops.	
5. Test V .....	29
Vocal reproduction of the notes $c^1$ and $a^1$ using a Seashore tonoscope.	
6. Test VI.	
Practice in the identification of the 88 piano notes.	
a. Group 2 .....	29
b. Retesting of members of Group 1 and 2 ....	33
c. Group 3 (selected music students) .....	33
7. Test VII .....	34
Identification of chords.	
a. Practice series.	
b. Single test.	
8. Test VIII .....	34
Identification of violin notes, with the movements of the observers recorded on a kymograph drum.	

9. Test IX .....	34
Series of identifications of tuning forks.	
CHAPTER IV. Data and discussion.	
1. Graphs and tables for the experiments with the piano .....	36
a. Average error frequency .....	36
b. Error frequencies for first and last trials ..	37
c. Average frequency of correct notes .....	38
d. Frequencies of correct notes in first and last trials .....	39
e. Frequency of interval errors .....	40
f. Average time per note (5 observers) .....	42
g. Average time for correct and incorrect notes (5 observers) .....	43
h. Individual practice curves (5 observers) ..	44
2. Correlations for the piano experiments between	
a. Initial and final performances.	
1. Based on errors .....	38
2. Based on correct notes .....	40
b. Number of correct notes and average error	38
c. Accuracy of judgment and size of semitone interval in vibrations .....	50
d. Average time and average error .....	43
e. Time per note and size of error .....	43
f. Practice and improvement .....	45
g. Musical training and improvement .....	45
h. Number of years of musical training and average error .....	45
i. Pitch discrimination and average error ....	46
j. Size of error at last trial and at retesting ..	46
3. Additional data.	
a. Average error in the piano test one year later	46
b. Individual frequency of correctness for the 88 notes on the piano .....	47
c. Average error per note, range of positive and negative errors and median position for each 3-semitone division of the 87 piano notes ...	50
d. Average error and average time in the pipe-organ tests .....	51
e. Violin errors and correct notes .....	52
f. Kymograph records .....	53



14

g. Results from the experiments with chords ..	54
h. Results from the experiment with tuning forks .....	54
i. Tonoscope records .....	55

CHAPTER V. Analysis of introspective reports.

1. Responses to the questionnaire .....	68
a. Musical training.	
b. Musical appreciation.	
c. Musical memory.	
d. Imagery.	
e. Methods for designating notes.	
f. Setting of notes.	
g. Aids and hindrances.	
2. Review of the introspections given at the end of each sitting .....	71
3. Abstracts from the introspections of the observes in Group 1 .....	72
4. Abstracts from the introspections given after the retesting on the piano one year later, the experiments with chords and the experiment with the tuning forks .....	82

CHAPTER VI. Summary and Conclusions.

1. The effects of practice .....	86
a. During the learning period.	
(1) The reduction of errors.	
(2) The increase in the number of correct notes.	
b. The following year.	
2. Group characteristics .....	86
3. Individual differences (Group 1) .....	87
4. Effects of timbre .....	88
5. Kinaesthesia during the judging period .....	88
6. Reproduction of notes .....	89
7. Discussion of possible criteria for tonal judgments .....	89
a. Pitch.	
b. Brightness and volume.	
c. Octave differences and vowel qualities.	
d. Octave similarities.	
e. Complexity of sound wave.	
f. Frequency and practice.	

## PREFACE

It gives me pleasure to express here my gratitude to the many people who aided in the progress of this investigation. John W. Baird, under whose instruction I first became interested in the problem, directed the initial stages of the work. He also brought me into communication with Professor C. A. Ruckmich who generously contributed the report of a similar experiment conducted by a student in his laboratory. Professor C. A. Seashore in correspondence suggested new aspects of the problem and methods of approach. I am especially indebted to Professor R. S. Woodworth for his constant encouragement and criticism throughout the investigation.

I wish to thank my fellow students, Miss R. S. Clark and Miss H. J. Sleeper, my sister, Mrs. E. S. Boegehold, and the many undergraduates at Smith College, who cooperated so heartily in this study. I also extend my thanks to the Department of Music at Smith College for affording the necessary facilities for carrying on the experiments.

# The Effects of Practice on Judgments of Absolute Pitch

## I

### INTRODUCTION

Many striking cases of absolute pitch memory have already been investigated, yielding interesting and valuable results. Individuals who possess this ability in a high degree are, however, unable to analyze their means of recognition and trace its development. They identify a particular note as  $c^1$  in the same way that they recognize any familiar percept or respond to any elementary experience. This attitude has led to the belief that certain highly favored people are endowed by nature with an unusual tone-recognizing mechanism which has been denied to others.

The present study has undertaken to investigate to what extent the average individual is able to identify notes, and to what degree his initial ability may be improved by training. By an analysis of the introspective reports given at these tests and a comparison of these reports with the objective results, it is possible that some light may be thrown on the general problem of judgments of absolute tone.

*Historical Survey* A review of the literature on this problem brings out many casual observations of the ability, many speculations as to its nature and its relation to other abilities and to practice effects, and also a number of systematic studies in this field.

Stumpf<sup>1</sup> writes as early as 1883 that he believes the individual differences in memory for pitch to be due to such variable factors as practice effects, general retentiveness and a certain inexplainable individual coefficient.

In tests on a group of musicians he finds that the different regions of the keyboard have varying degrees of difficulty, and that clang-tint affects the ability to identify notes to a great degree. Different observers require different lengths of time for judgment, and the same observer varies according

---

<sup>1</sup>C. Stumpf, *Tonpsychologie*, 1883, pp. 286-313.

to the region to be judged. Stumpf's own results from six tests on the piano are as follows:<sup>2</sup>

Tests	C <sub>1</sub> to B <sub>1</sub>		g to eb <sup>3</sup>		f# <sup>3</sup> to a <sup>4</sup>	
	Number Correct	Number Presented	Number Correct	Number Presented	Number Correct	Number Presented
1	1	7	10	10	3	13
2	3	14	9	9	3	13
			a to	g# <sup>1</sup>	g <sup>2</sup> to	f# <sup>4</sup>
3	3	13	12	13	6	13
4	4	13	12	13	6	13
5	3	13	10	13	5	13
6	8	13	13	13	6	13

His errors are mostly intervals of a semitone or a whole tone. A few extend to the tritone, but never include fifths or sixths. Little account is taken of octave errors, because musicians make slight use of the symbols designating the octave and consequently do not easily differentiate them. The effects of practice are noticeable in these results.

Stumpf reports a test he made on a girl eight years old, using the sixty piano notes between C and c<sup>4</sup>.<sup>3</sup> Thirty-nine notes were correctly named. Most of her errors lay in the three-accented octave and occurred among the accidentals. Black keys were confused with other black keys, as c#<sup>3</sup> with f#<sup>3</sup>. She was less able to identify chords than single notes.

He reports the case of a man who could recognize the keys of chords played on the piano or played by an orchestra, but not played on an organ.<sup>4</sup> And he was unable to identify individual notes. Stumpf explains this on the basis of the direction of musical interest.

Several people are recorded who possessed interval memories inferior to their memories for absolute pitch.<sup>5</sup> If a piece was played in an unfamiliar key, even though it might be only a semitone or a whole tone higher or lower than the accustomed key, it was necessary for them to transpose individually each note to the new key before being able to sing or play the melody.

Further conclusions from Stumpf's investigation of the subject are that absolute pitch memory appears very seldom among women, and that it is not an indispensable factor in

<sup>2</sup>*Ibid.*, p. 310.

<sup>3</sup>C. Stumpf, *Tonpsychologie*, II, 1890, pp. 553 f.

<sup>4</sup>*Ibid.*, p. 554.

<sup>5</sup>*Ibid.*, p. 555.



musical ability.<sup>6</sup> He attributes much importance to the element of frequency in establishing a good memory for pitch, but regards as more fundamental the individual's initial discriminative ability and his inherent retentiveness for certain kinds of auditory impressions.

In 1892 von Kries<sup>7</sup> cites a number of observations which he made on himself and on a few others who possessed a good memory for absolute pitch. Violin pieces are occasionally written for violins to be tuned a semitone higher than normally. Some violinists find it impossible to play music written in this way, others experience no difficulty. Hence von Kries conjectures that the latter do not have absolute pitch memory. He agrees with Stumpf in finding the middle range the easiest and the surest to judge. The notes above  $c^4$  and below  $C$  are difficult to hear distinctly.

Arranging the different musical instruments in order of ease of judgment, the piano comes first, then the greater number of string and wind instruments, then the voice, tuning forks and whistles. Since the different clang combinations influence greatly the ability to judge pitch, he concludes that memory images are not the essential condition,—in other words, that absolute pitch memory does not involve a comparison judgment. He explains this form of memory from the Lehmann association point of view. It is based on a kind of limited association, which holds only for certain clangs. He finds in studying the different clangs that recognition may have all grades of clearness. A sharp line is not to be drawn between the recognizable and the non-recognizable tones. With long practice he was able to reduce the limits of the unrecognizable, but was never able to pass from indirect to direct judgments.

He rejects the theory of muscular accompaniment, for in his own case he has more success in judging the pitch of soprano voices than male voices, and tones which he himself cannot sing at all. Frequency of hearing certain notes does not fully explain results for him, for in spite of his wide experience with male voices and his slight experience with the violin, judgments of pitch in the first case are poor and in the second good. He hears one-part singing oftenest yet can

---

<sup>6</sup>Stumpf, *loc. cit.*, I., p. 286.

<sup>7</sup>J. von Kries, Ueber das absolute Gehör. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 1892, 3, pp. 257-279.

judge the pitch of several-part singing more easily. He finds many cases of people who can recognize keys and chords, but not single notes.

Richness in overtones seems to be the most predominant factor contributing toward absolute pitch memory, although these tones need not be present consciously to the observer. The fact that the human voice gives difficulty he thinks may possibly be explained by the interference of the vowel sounds in singing.

A report is made by Meyer<sup>s</sup> in 1889 of an investigation in which he and one other observer attempted to develop absolute pitch memories through practice. They worked two months using the notes produced by sixteen tuning forks varying from 100 v. d. to 4000 v. d., and four months using from ten to thirty-nine notes produced on the piano. More than half their judgments were correct and the size of the errors was very small. After discontinuing practice for several years Meyer found they had lost the greater part of what they had acquired.

The following questionnaire was sent to one hundred musicians by Abraham,<sup>o</sup> who bases his conclusions regarding memory for absolute pitch in part upon the replies he received:

1. How long have you possessed absolute pitch memory?
2. Do you play or sing? How long?
3. Do you compose? Do you improvise on some instrument? Is it difficult for you to find the correct bass for the melodies you know?
4. Does your memory for pitch include the ability to produce by singing or whistling certain notes as well as the ability to name correctly notes you hear?
5. How do you arrive at your correct judgment? Is it immediate judgment by a conscious comparison with a note in memory or by singing the note heard? Can you think of a note without singing it or hearing it? What timbre has this note? Do you compare the note heard with the relatively lowest or highest singing tone which you can produce?
6. Is it easier for you to recognize a note in a melody or the key of a chord than a single note?

---

<sup>s</sup>M. Meyer, Is the Memory of Absolute Pitch Capable of Development by Training? *Psychological Review*, 1899, 6, pp. 514-516.

<sup>o</sup>O. Abraham, Das absolute Tonbewusstsein. *Sammelhefte der internationalen Musikgesellschaft*, 1901, 3, pp. 1-86.

7. In singing correctly a note from memory do you image its notation on the staff, or do you think of a song which begins on this note?
8. Has your pitch memory upper or lower limits in range? Do you judge one octave better than another?
9. How loudly and how long must the note be sounded for you to judge it easily?
10. What errors in judgment do you sometimes make? Octave, fifth, semitone, etc.?
11. Have all the *as* for example, some likeness which distinguishes them from *b*, *c*, etc.? Is this similarity present in a lesser degree between *a* and *e*?
12. Can you tell when an instrument is a quarter or an eighth of a tone higher or lower than another if a long period of time intervenes between the two sounds?
13. When a song is transposed by an accompanist can you sing it easily, or must you transpose it yourself mentally?
14. Does difference in timbre affect your ability to judge notes?
15. Have you a good memory for melodies? Must you think of a melody always in its correct key, or can you imagine it in any key?
16. Did you develop or improve your memory for pitch by practice? Is it hereditary?
17. Do you think of colors in connection with notes?

Abraham also made experiments in regard to the influence of pitch, intensity, duration of sound, time-interval between notes and tone-color on memory for absolute pitch. He found that the range of the ability is wide and variable, and does not coincide with the range of ability for sensible discrimination. Loud notes in which the high overtones are prominent are judged higher than they are. The pitch judgment does not require a longer stimulus time than that necessary for perception, but repetition of the stimulus increases the assurance. Notes without overtones are judged lower than they actually are.

According to Abraham's results some people possess absolute pitch memory without special practice, others are able to acquire it through practice, and others are never able by any means to acquire it. He attributes this ability to an individual factor which may be due to a certain fineness of sensory apparatus by which individuals recognize separate tone



qualities in their finer differences, or to certain unique cerebral conditions. This memory depends on associations formed in the auditory realm alone, and therefore lacks the many aids to memory which are found in other types of learning. "If a child is shown a clock, he sees it, feels its shape, hears it tick, feels the coldness of it, etc. The different senses help him to realize the object, so that he can always recognize it as a clock in the future . . . . Absolute color consciousness seems to be more common than absolute tone consciousness. Very fine differences of color are recognized, and color shades are produced in imagination by the name. However, I find that the color always attaches itself to some definite object. I cannot imagine brownish red purely as a color, but must think of the portière in my room. We cannot use instances like this to help us in our decision of pitch. We can only fix the pitch by itself in our memories."<sup>10</sup> Other factors which hinder the development of absolute pitch memory are the variations in the keys in which a child hears the same song sung and played, the variations in the standard pitch for different instruments for different purposes and the use of the tonic sol fa system, where *do* may sometimes be *c*, sometimes *e<sub>b</sub>*, etc.

Certain indirect methods may aid in determining the correct pitch of a note. The pain sensations which accompany very high notes, 'colored-hearing,' feelings of tension in humming a note, emotional reaction, and the characteristic pitch of the ear drum might be used as bases for inference. Visual images of the note or the key of the piano, or of oneself sitting at the piano, or the scene from some opera are sometimes used. These devices have no part in a good tonal memory, but they may lead to a firm impression of the note on the memory.

A person who has the "power of true tone imagination," or the ability to call up a note at will, is able to make exceedingly fine judgments of absolute tone, such as judgments of notes less than a quarter of a tone apart, because he can compare each note he hears with the note in his imagination. These two aspects of tonal memory are inseparable, but each marks a different stage in the development of the ability.

Absolute key consciousness is often confused with the recognition of certain physical differences in the notes of different

<sup>10</sup>Abraham, *loc. cit.*, p. 9.



instruments. The white keys on the piano give a louder sound than the black keys, partly because one uses a more powerful touch on pressing them down, and partly because the subduing effect of the hammer, as the white keys are played more frequently, is very soon lessened disproportionately. The open strings of bowed instruments sound brighter than the other notes, so that they give the keys in which they are used a great deal, especially those in which they are the key notes, a brighter character. The natural tones of wind instruments sound brighter than those made by stops, etc. One's feeling for keys is undoubtedly affected by associations of various kinds. If a composer has written a powerful piece in *d*-minor, later musicians come unconsciously to associate that key with power.

Experiments made by Abraham on a man who claimed to be able to recognize keys but not single notes, show that only *f* $\sharp$ -minor could be correctly recognized, and only when played on his own piano. Abraham concludes that the existence of absolute key consciousness without absolute tone consciousness is very uncertain.

Abraham believes this ability is an important part of a musician's equipment, but taken alone need not indicate musicianship. He finds it frequently accompanies imperfect melody memory, which he explains on the theory that the association paths within the auditory area from tone to tone are more numerous or more closely knit together than the paths from tones to other sensory areas, as to a word, a visual image, etc.

Other possible explanations for the decidedly different character of different keys on piano and bowed instruments are offered by Helmholtz.<sup>11</sup> "*C*-major and the adjacent *d* $\flat$ -major have different effects. That this difference is not caused by difference of absolute pitch can be determined by comparing two different instruments tuned to different pitches. The *d* $\flat$  of the one instrument may be as high (*sic*) as the *c* of the other, and yet on both the *c*-major retains its brighter and stronger character, and the *d* $\flat$  its soft and veiled harmonious effect."<sup>12</sup> This may be due, as Abraham has said to the method of striking the short black keys of the piano, or as

<sup>11</sup>H. von Helmholtz, *Sensations of tone*, Trans. by A. J. Ellis, 1895, pp. 310-330.

<sup>12</sup>*Ibid*, p. 311.

Ellis<sup>13</sup> suggests to the different leverage of the black keys, so that each key has a different distribution of the stronger and gentler quality of tone among the varying degrees of the scale. Further the difference made in the tuning of those fifths which the tuner keeps to the last, and on which are crowded the whole of the errors in tuning the other fifths in the circle of fifths, may be regular and contribute to this effect.

On bowed instruments the notes from the open strings color the key by the inequality of intonation. The open strings give perfect fifths, but in playing in different keys if each note has the same sound throughout, the other fifths will not be perfect. Hence the scales of the various keys will differ in intonation and in character. Wind instruments illustrate this in an even more striking manner.

Another possibility is that  $g^4$ , the proper tone of the human ear, which sounds peculiarly shrill at all times, may color those notes which possess it as an upper partial, as  $g^3$ ,  $c^3$ ,  $g^2$ , and give to them a brighter and more piercing character than other notes.

Boggs<sup>14</sup> investigated memory for absolute pitch in a group of especially endowed individuals, all of whom had a musical inheritance and early musical training. Her observers displayed immediateness of recognition and sureness of judgment. Concentration of attention was found to be of great importance. Extraordinarily good discrimination of pitch was not observed and imagery did not seem to be particularly necessary. A strong emotional feeling for music as a whole was shown, but not for isolated notes. Her observers reported that they were conscious of overtones and that the pitch of the particular piano used made practically no difference in their judgments.

This factor of attention to overtones led Boggs to believe that each note is heard in its relationship to other notes as part of a tonal system. This view is supported by the fact that chords are frequently more easy to identify than single notes. This tonal system is a qualitative one comparable to a phase of the spectrum in vision, each system being regulated in part by differences in vibration rate. The tonal scale corresponds to the brightness variations. Hence the average

---

<sup>13</sup>*Ibid*, p. 311.

<sup>14</sup>L. P. Boggs, Studies in Absolute Pitch. *American Journal of Psychology*, 1907, 18, pp. 194-205.

person might hear the scale in the same way that a color-blind person sees the spectrum as graduated succession of brightnesses.

Writing on tests for musical ability, Rupp<sup>15</sup> states that memory for absolute pitch is in general indicative of musical ability. Although it does not constitute the essence of this ability, it is closely bound up with other musical qualities. Since he examined only a small number of cases, he cannot attach a high degree of certainty to his correlation.

He suggests the method of minimal changes as advantageous for the study of the upper and lower thresholds for the memory of notes. For instance by slowly changing the pitch on a variator, one can discover within what limits the sounds seem like  $a^1$ ,  $c^1$ , etc.

He finds, as has been before mentioned, that the sureness and exactness of memory is not the same for different notes and octaves. More mistakes were made in the extremes than in the middle octaves.  $A$  and  $c$  in the different octaves were recognized more quickly than other notes which were named with hesitation and sometimes by reference to  $c$  and  $a$ . The differences in the tone color of different instruments was a hindrance which led to indirect judgments such as comparing the tone of new timbre with more familiar tones. By continued practice Rupp believes it is possible to reach correct absolute pitch judgments without intermediary aids. A long series of impressions of notes would lead in time to a firm, lasting impression. This may be either acoustic or kinaesthetic memory.

In reproducing notes as distinguished from recognizing notes heard, the individual's technique in handling his instrument is an important factor in his success. Hence a good vocalist without absolute pitch memory might make a better showing than a poor singer with memory for pitch.

The results from experiments on absolute pitch memory, conducted by a student in the Psychological Laboratory at the University of Illinois, were made available for the writer through the kindness of Professor C. A. Ruckmich.<sup>16</sup> In order to investigate the problem, "Can absolute pitch be cultivated," practice experiments were made individually on two musical

---

<sup>15</sup>H. Rupp, Ueber die Prüfung musikalischer Fähigkeiten, Teil I, *Zeitschrift für angewandte Psychologie*, 1914, 9, pp. 1-76.

<sup>16</sup>Can Absolute Pitch be Cultivated? Manuscript, 1916.



adults who possessed varying degrees of accuracy in pitch memory, two children, nine and twelve years of age, and the students in an ear training class. The observers began by trying to impress on their memories the notes  $c^1$ ,  $e^1$  and  $g^1$ . After a time all the white keys in the once accented octave were undertaken and later the range was extended to two octaves. The black keys in the middle range were then added. The class in ear training was given notes taken at random from the entire keyboard.

The results from this practice evidence a degree of improvement. Both the size and the percentage of error are decreased. The average time from all the experiments for the correct judgments is slightly less than that for the incorrect, although this does not always hold in single tests. The writer concludes that accuracy of judgment varies directly with the speed.

In testing the children the experimenter found it impossible in spite of his efforts to lead them to think of a note in terms of absolute pitch. They insisted on making their judgments relative to some other note, usually  $c$ , which they seemed to remember remarkably well.

On account of the difficulties involved in explaining the many aspects of memory for absolute pitch as due either to the complexity of the sound wave, or to the familiarity with certain clangs, Köhler<sup>17</sup> is led to believe that recognition of notes is in many cases the result not of a memory for pitch, but a memory for some other feature of tonal body. He quotes such aspects as the immediateness of the recognition of those who possess good tonal memory, their proverbially poor interval memories, their incompetence in the face of some new clang tint, the superiority of the recognition of chords over the recognition of single tones, and the frequency of errors of the octave interval, of fifths and of fourths.

With the importance of tonal body in mind, Köhler made an experimental study on himself trying to learn the notes of the white keys between  $c$  and  $b^3$ . After fourteen days practice he was convinced of the value of this method. Half his 220 judgments were correct and the majority of his interval errors corresponded to the errors of individuals in whom ab-

---

<sup>17</sup>W. Köhler, *Akustische Untersuchungen, Zeitschrift für Psychologie*, 1915, 72, pp. 159-177.

solute pitch memory was native, i. e., errors of the octave, fifth and fourth.

His results lead him to the conclusion that memory for tonal body may be the criterion for correct tonal judgments and affords a plausible explanation for many of the striking characteristics alleged to belong to memory for pitch. Neither pitch nor tone body alone cover all cases, therefore either one may be effective in the judgment. He explains the difficulty in judging tones from human voices as the result of the continual shifting of the register by the variation in vowel sounds. He quotes Max Meyer and Heyfelder as supporting his view that pitch alone is not adequate to explain the problem.

The most extensive investigation of the nature of absolute pitch memory was made by Baird.<sup>18</sup> All his observers were more or less highly trained musicians who possessed some degree of memory for absolute pitch. Nine observers took part in identifying the pitch of the eighty-eight notes of the piano, taken in random order (i. e., naming the note definitely as  $c^3$ ,  $A_2$ , etc.), and five observers took part in identifying sixty notes of the pipe-organ, using the flute, diapason, reed and string stops, fifty-five notes ranging from  $B_b$  to  $g\sharp^2$  sung by four vocalists, soprano, contralto, tenor and bass, twenty-four notes,  $c^1$  to  $c^3$ , of the flute and the clarinet, fourteen tuning forks ranging from  $c$  to  $c^3$ , twenty-three notes beyond the keyboard of the piano from  $d^5$  to  $c^8$ , using König bars and the Galton whistle. In another group of experiments he obtained the reaction time by means of a Hipp Chronoscope. The observer was asked to name notes of the piano and organ without designating the octave. In a third group of experiments six observers were asked to reproduce the notes of an octave by means of the Tonvariator. Determinations were made in an ascending and a descending direction.

His results in regard to note errors show great individual variation. (1) Although the observers belong to a highly selected group, their average error in identifying piano notes varies from 1.1 per cent. to 74 per cent. This is true also for the identification of notes of other clang tints. (2) A variation of accuracy with clang tint is found in the case of each observer. Arranging these in order of the ease of identification for each person, marked individual preferences are found.

<sup>18</sup>J. W. Baird, Memory for Absolute Pitch, *Studies in Psychology, Titchener Commemorative Volume*, 1917, pp. 43-78.

On the whole the piano notes are most easily identified. Then follow in approximate order the notes from the organ, the flute, the clarinet, and the voice. The results from the other clang tints are uncertain. (3) The middle range of the piano is in every case most accurately named. The average percentages of error distributed over the tonal scale by octaves are for nine observers as follows:

A <sub>2</sub> -B <sub>2</sub>	C <sub>2</sub> -B <sub>1</sub>	C-B	c-b	c <sup>1</sup> -b <sup>1</sup>	c <sup>2</sup> -b <sup>2</sup>	c <sup>3</sup> -b <sup>3</sup>	c <sup>4</sup> -c <sup>5</sup>
57.1	51.7	42.0	28.5	19.9	23.5	30.8	43.4

Greater individual variation occurs in identifying pipe-organ notes. (4) The naturals are usually identified with greater ease than the accidentals, although here again great individual variation is found. Every observer was more successful in naming certain notes of the octave than others in the case of both piano and organ notes with little coincidence of preference in the same individual with these two instruments. The averages from all observers show that *f* and *c* are most accurately identified on the piano and *g* and *c* on the organ, whereas *c*<sup>#</sup> and *b*<sup>b</sup> are least accurately identified in both cases. (5) A central tendency of judgment is noticeable. Low pitched notes tend to be judged too high and high pitched notes too low. Overestimations of pitch are relatively more frequent for piano notes and underestimations for organ notes. (6) The size of the error varies from one semitone to two octaves. The proportion of gross errors increases with the decrease in number of correct judgments from one person to another and with the same persons from one instrument to another. (7) Certain observers habitually make the same mistakes. Errors of a fifth occur in 51 cases, errors of a fourth in 54 cases, errors of a sixth in 70 cases, and errors of a third in 80 cases. (Baird's observers show a tendency to call many of the uncertain notes *g*, for *c*, *d*, *e*, *f*, *b*<sup>b</sup> and *b* are wrongly identified as *g* a large number of times.) From these findings Baird concludes that the notes least subject to tonal fusion are most subject to confusion. Although the observers believed they were aware of qualitative differences between naturals and accidentals, no evidence is found to support their view. In the case of errors among the accidentals, 38% are misidentified as naturals, 12.1% as accidentals. On the other hand in the case of errors among the naturals, 20.1% are misidentified as accidentals, and 29.8% as naturals. This shows that the response of 'natural' occurs more than twice as frequently



as the response of 'accidental.' (8) Some slight ability to identify pitches beyond the range of the piano is found. The observers report that above *c'* the pitch is indistinct and smothered by noise.

Octave errors are common in the identifications of all the observers in every kind of clang tint, but are subject to much individual variation. They are more frequent in the more accurate observers with one notable exception, and more frequent with relatively easy clang tints. Baird distinguishes two stages in the process of identification, an initial stage which consists in a prompt naming of the note and a subsequent stage in which the observer names the octave to which the note belongs. These two responses are usually separated by an appreciable interval, and the degree of subjective assurance is less for the octave identification than for the name of the note.

The results for identification time show that piano notes are more quickly recognized than organ notes, that the notes in the middle range are more quickly recognized than those at the extremes, and that a fairly close correspondence exists between accuracy and promptness of response.

Wide individual differences are shown in the ability to produce given notes by means of the variator. The degree of accuracy of each observer corresponds fairly well with her degree of accuracy of identification of notes. Wide tonal "bands" with considerable overlapping of neighboring notes satisfy those who are least accurate in their identifications. Those most accurate in the previous tests reproduce notes with relatively narrow bands, without overlapping, located at appropriate regions in the tonal scale.

Baird rejects as inadequate the theory that complexity of sound wave is the important factor in the identification of pitch. He advances an alternative explanation which is supported by results of investigations on vocality.<sup>19</sup> This states that distinctive qualities attach to different regions of the

---

<sup>19</sup>Köhler found in using tuning forks as stimuli that a series of tones ranging from 163 v. d. to 4000 v. d. gives a series of vowel qualities which can be arranged like those of color. As red shades gradually into orange and orange into yellow, similarly in the tonal system the vowel quality *u* (moon) passes through *u-o*, *o-u* to *o*, and so on to *a* (father), *e* (prey), and *i* (machine). Later in using Stern variators with interference apparatus he obtained the pitches of the pure vowels and added *e*, *f* and *ch* to the upper range and *m* to the lower. His pure vowels fell regularly into octaves.

tonal scale. One may experience these qualities a single time possibly and from that time be able to recognize them when they recur just as the normal individual recognizes the elementary experiences after a single occurrence. On examining the process of recognition in his observers Baird does not find an initial learning period filled with a variety of mental content, which later fades out as recognition becomes immediate, such as is necessary for acquisitions of mental or motor skill. His observers claimed that they had never attempted training of this kind, but had simply discovered at an early age this ability, which apparently remained unchanged. The literature on practice experiments in this field strengthens his belief that "memory for absolute pitch is based upon an ability to detect the presence of the *c*-quality which is obscurely present in every *c*, of the *d*-quality which is obscurely present in every *d*, etc. . . . Given this base as the advent of the capacity to recognize pitches is inevitable and abrupt."<sup>20</sup>

The importance in pitch recognition of the so-called tonal attributes of quality and vocality, defined by Révész<sup>21</sup> as that which recurs at every octave and that which is different in different octaves, has been emphasized by a number of writers. Max Meyer,<sup>22</sup> writing in 1914, believes that the distinction between tonality and vocality is essential for an understanding of memory for absolute pitch. Révész<sup>23</sup> states that there may be several kinds of memories for absolute pitch, in which one or the other or both attributes play a part. A musical interval can be either a span in vocality, or a relation of tonality, or both.

The psychological status of these two characteristics is a matter of dispute. Do they rank as attributes of sensation, or as perceptions, or are they completely outside the tonal series? Stumpf<sup>24</sup> denies on the basis of introspection that vowel qualities are comparable to the color series, and explains octave similarity as due to the degree of fusion, an ultimate fact of hearing. In 1914 he writes that the primary

---

<sup>20</sup>Baird, *loc. cit.*, p. 78.

<sup>21</sup>M. Meyer, Review of Révész, *Psychological Bulletin*, 1914, 11, p. 349.

<sup>22</sup>*Ibid.*

<sup>23</sup>*Ibid.*

<sup>24</sup>Quoted from G. T. Rich, A Study of Tonal Attributes, *American Journal of Psychology*, 1919, 30, pp. 121, 125.



qualities lie within a single octave.<sup>25</sup> Rich<sup>26</sup> reports a recent experimental study of tonal attributes in which he concludes that vowel quality judgments are perceptions, but he is in doubt whether to regard tonality as attributive or perceptual. Watt<sup>27</sup> denies the view that vowels are primary qualities of hearing or in any sense attributive. He believes vowels hold a mid-position between tones and noises. Also octave similarities do not require a theory of octave qualities in explanation but can be interpreted as the result of the symmetry of volumic relations.

Watt<sup>28</sup> offers an explanation for absolute pitch memory on the basis of his analysis of tones. Tones are systems of sound possessing six attributes, (1) intensity, (2) extension or volume, (3) systemic order, under which heading pitch falls, (4) a single quality characteristic of all tones, (5) temporal order, and (6) durance. Pitch is not to be identified with tone, but is to be regarded as the intensively predominant order of the whole volume, which is symmetrically balanced in relation to pitch. Differences in pitch illustrate the ordinal series in sound as distinguished from a qualitative series. A unit in the tonal series is better described in quantitative terms as 'this one,' and 'that one' and 'the next one,' than in qualitative terms as 'this sort of one,' 'that sort of one,' and 'the other sort of one.' A scale is a series of volumes of constant proportions, an octave simply a pattern of volumic relations, which has set an artificial standard for the whole range, and a musical interval an intensive volumic outline. Both absolute ear and relative ear are based on these relationships. For the former a point of reference in auditory orders is required. "Absolute ear emerges when the natural absoluteness of tonal orders maintains its efficiency in spite of the tremendous emphasis laid on relativity or proportion in music."<sup>29</sup>

Naming tones correctly is on a par with localizing accurately a point touched on the skin. Each is dependent on the attribute of order. The chief reason why many fall short in the first case is due to the extraordinary emphasis laid by

---

<sup>25</sup>R. M. Ogden, Tonal Attributes, *Psychological Bulletin*, 1917, 14, pp. 161-162.

<sup>26</sup>G. T. Rich, A Study of Tonal Attributes, *American Journal of Psychology*, 1919, 30, pp. 121-164.

<sup>27</sup>H. T. Watt, *The Psychology of Sound*, 1917, pp. 41-53.

<sup>28</sup>*Ibid.*, pp. 15-53.

<sup>29</sup>*Ibid.*, p. 200.

music on the relationship of pitches. The variation in absolute pitch of musical instruments is in part responsible for the defect. Some people may have a special refinement of hearing, such as a greater delicacy of volumic outline and especially of predominance than others have. Or a highly favored auditory disposition might give them the power to maintain their absoluteness of ear in spite of the universality of musical relativity. In this case everyone naturally possesses absolute ear, and later loses it, or loses the power to convert it into absolute nomenclature.<sup>30</sup> Octave errors in absolute ear are the inevitable results of the relativity which the octave brings. Even a well-maintained absolute ear is in part subdued to musical relativity, regarding as of first importance the placing of the pitch in the octave, and of minor importance accuracy in specification of the exact octave.

Myers and Valentine<sup>31</sup> made a study of the different attitudes which different people assume toward tones. They analyzed four possible attitudes (1) the intra-subjective attitude, which involves certain physiological sub-aspects in which the observers characterize the note as 'strident' or 'piercing' and conative subaspects such as impulses to motion and efforts to determine meaning, pitch or interval; (2) the objective attitude, in which the observer relates the sound to some standard of purity, pitch, etc.; (3) the character attitude in which the observer tends to personify the tones heard; and (4) the association attitude including musical (fused) associations and other than musical (non-fused) associations. Their results for single tones are as follows:

	Intra-subjective Attitude	Objective Attitude	Character Attitude	Association Attitude
Highly Musical Observers	28%	39%	15%	18%
Less Musical Observers	22%	32%	13%	33%

The attitude which an individual assumes when listening to isolated notes will affect the possibility of his regarding

<sup>30</sup>E. M. von Hornbostel (Ueber vergleichende akustische und musik-psychologische Untersuchungen, *Zeitschrift für Angewante Psychologie*, 1910, 3, pp. 465ff.) states that children can be trained to develop absolute pitch memories, and that the less contact they have had with music, the better this training. Parrots and starlings have been observed to reproduce melodies only in the original keys in which they were sung.

<sup>31</sup>C. S. Meyers and C. W. Valentine, Individual Differences in Attitude toward tones. *British Journal of Psychology*, 1914, 7, pp. 68-111.

tones as part of an ordinal series with absolute points of reference. These results show that the highly musical observers are more apt to have the analytical point of view for single tones than the less musical observers.

Summarizing the results of these writers, we find many points of similarity and of difference. The reason why the ability is so restricted is believed according to Abraham to depend on the facts (1) that absolute pitch forms its associations in only one sense realm, the auditory, (2) that songs are transposed freely from one key to another, (3) that the standard pitch of instruments varies greatly, and (4) that the use of the tonic sol fa system emphasizes the side of relativity in music. Watt also stresses this last factor of relativity in musical training.

Various characteristics and accompaniments of the ability are: (1) the middle range is universally regarded as the easiest to judge and gives the surest judgments. Baird finds *f* and *c* on the piano most frequently judged correctly and *g* and *c* on the organ. (2) Clang tint directly affects judgments. Arranging the instruments in order of ease of identification Baird puts the piano first, then the organ, the flute, the clarinet and the voice. Von Kries places tuning forks and whistles after the voice. (3) Stumpf found that the speed of judging time varies from person to person and with the same person from one clang to another. Ruckmich's student adds that speed and accuracy vary together. (4) Stumpf's errors in identification were mostly semitone and whole tone errors diminishing gradually in number as their size increased. No errors of the fifth or sixth occurred. Baird's errors ranged from a semitone to two octaves. He found a large number of octave errors, 70 errors of a sixth, 51 errors of a fifth, 54 errors of a fourth, and 80 errors of a third. Köhler emphasizes the predominance of errors of a fifth, a fourth and an octave amongst observers who possess genuine absolute pitch memory. (5) Stumpf found that errors of black keys were due to confusions with other black keys. Baird's results contradict this. (6) Abraham suggests that poor melody memory is apt to accompany memory for absolute pitch.

Chords are frequently claimed to be easier to identify than single notes. One of Stumpf's observers could name single notes correctly but not chords, one could name chords but not single notes. Abraham's results led him to feel doubtful of



the existence of absolute key consciousness without absolute tone consciousness.

Baird found the ability to reproduce given notes to run fairly parallel with the ability to name notes heard. Abraham thinks the two abilities need not necessarily be at the same level of development, but that both represent the same general capacity.

Many possible aids in making absolute pitch judgments are enumerated by Helmholtz and Abraham. They are (1) feelings of tension in humming, (2) emotional reaction, (3) pain sensations caused by high notes, effects of the characteristic pitch of the ear drum on  $g^4$  and on notes which possess  $g^4$  as one of their overtones, and (4) the quality or intensity of tone as a result of the method of striking the black and white keys. In judging keys on the piano the effects of constant errors in tuning, and in judging keys on bowed and wind instruments the variation in intonation of the fifths might serve as secondary criteria.

The essential criteria for absolute pitch memories vary from writer to writer. Writers are fairly agreed that the muscular accompaniment cannot be wholly responsible for these judgments. Overtones or complexities of tone are rejected by Baird, and are held to be important by von Kries and Köhler, who explain the discrepancy found for singing tones as due to the influence of vowel sounds which continually shift the register of the voice, and also by Boggs, who believes that they lead us to a knowledge of the extended tonal system, comparable in many ways to the spectral series. The place of frequency and of practice in establishing memory for pitch is much disputed. Stumpf finds it important, provided other criteria are present. Von Kries thought it reduced the limits of the unrecognizable, but never to the extent of converting an indirect judgment into a direct judgment. In Meyer's experiments practice developed a temporary ability which soon was lost through disuse. Abraham thinks practice may be beneficial with some people, wholly unnecessary for others and quite useless for still others. Baird denies that any evidence exists that training will originate the ability if it is not to some degree present already. On the other hand Rupp and Köhler agree that repetition will produce a firm lasting impression which according to Rupp may be either auditory or kinaesthetic.

The ultimate basis for absolute pitch judgments is finally (1) for Stumpf the inexplicable individual coefficient, (2) for Abraham either exceptional fineness of the sensory apparatus which enables an individual to recognize separate tone qualities in their finer differences, or certain unique cerebral conditions, (3) for Baird a discovery of the distinctive quality present in every *c*, in every *d*, etc., and (4) for Watt the attribute of order which is inseparable from every auditory sensation. This may be overshadowed by circumstances and is probably preserved either by an individual's greater delicacy of volumic outline or by a more highly favored auditory disposition.

## II

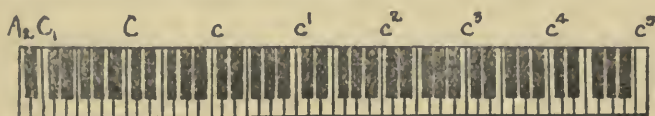
### OBSERVERS

Observers who were already trained to some degree in introspection were selected to take part in the experiments. The nine members of an advanced laboratory course in psychology and two instructors in the department formed the first group of observers in the practice series. Group 2 was made up of eighty students, members at that time of the required introductory course in psychology. The majority of these students were juniors, a small number were sophomores and seniors. They ranged musically from a few who excelled in violin or piano study in the department of music through varying degrees of ability to a few who had had no training of any kind or interest in music. Although this group, totally unselected, taken from a required course at the college, represents as widely distributed a sampling as is possible in a college community, it has, however, already been subject to much selection of entrance and cannot be taken as representing the average person. The training in analysis and introspection which the observers were obtaining in their psychology course was believed to be a valuable asset for this study. Group 3 contained three selected members, a junior and a sophomore of exceptional ability in music, who possessed an evident degree of absolute pitch memory, and the instructor in Ear Training and Harmony classes under the Department of Music, who did not possess memory for absolute pitches. During the summer tests were given individually to observer Bd, who had had considerable musical and psychological training.

### III

#### EXPERIMENTS

The members of Group 1 took part in the first series of practice experiments. The method of conducting these tests was as follows: The observer was seated with her back to the piano<sup>1</sup>, facing a cardboard piano keyboard which bore the symbols for the different octaves.



The experimenter seated at the piano held a list<sup>2</sup> of the 88 piano notes arranged in random order avoiding octave sequences and any sequence less than two octaves apart. The experimenter gave a signal, played a note mediumly loud and with a stop watch noted the time necessary for the observer to respond aloud with the name of the note. The note was repeated as often as the observer wished to hear it. As soon as the response and the time were recorded by the experimenter, the next note was sounded, continuing in this way until all 88 notes were given. A sitting lasted from about twenty minutes to forty minutes according to the speed of the observer.

Each of the ten members of Group 1 was given an individual preliminary trial with instructions only in regard to the distinctive symbols for each note. The notes were to be definitely located within their proper octaves. Before this trial the observers were allowed to play a few moments on the piano in order to gain a general idea of the quality of its tone. At the close of this first sitting the group was divided into two sections.

*Test I* Six of its members were instructed to practice the entire keyboard, using this piano always, for about forty minutes a week, in four ten minute periods, if possible, and to at-

<sup>1</sup>A new upright Knabe piano tuned relatively high ( $a^1 =$  about 445 v. d.) was used in all the different series of tests unless otherwise specified. The piano was located in a music practice room containing no hangings to soften its exceptionally clear resonant tone. During the year it was used only by the observers in these tests, and was tuned at frequent intervals.

<sup>2</sup>About fifteen of these lists were compiled during the course of the experiment.



tempt to devise some method by which they might be able to retain the pitches of the different notes. Memory for absolute pitch without reference to any secondary criteria was to be the goal, but any method might be used to reach this. Tests followed weekly the periods of practice throughout the greater part of the college year, 1917-1918, with some slight irregularities. At the end of each sitting introspections were written by the observer, especially regarding the method attempted in practice and its usability in the test. After a few weeks the writer acted as experimenter for all the other observers throughout the remainder of the investigation. Each of the other observers was at some time experimenter when the writer was observer. This brought about a fair uniformity of general procedure. The experimenter attempted always to employ a constant method of striking the notes producing equal degrees of intensity for the full range of notes. The duration of the notes differed for the different observers. Some asked for a number of quick repetitions, others preferred to have the tone ring out without interruption. If the response was immediate the sound was at once stopped; if there was hesitation, it was continued until the note was named.

*Test II* Four members of Group 1 undertook a somewhat different procedure. Instead of attempting to familiarize themselves with the entire keyboard at once, they began with a single octave, the once accented octave, and increased the range gradually, adding first the next higher octave, then the next lower octave after practicing for two weeks before each new addition. The presentation order of notes used by the experimenter was as varied as was possible, but at first, when the range was small, it was necessary to have an interval of fifteen seconds between the notes during which time the observer's attention to the previous note was distracted by unharmonious chord combinations extending over a wide range. Aside from this difference test II was conducted in the same manner as test I. This test continued through a part of the academic year 1917-1918.

*Test III* At the same time a test was begun in which the 31 notes on the violin, from  $a_1$  to  $d^3$ , were used instead of the 88 piano notes. One observer from Group 1 carried this through as a practice experiment for half the year, then later joined



the group in test I. The method of procedure was identical with that of test I. Another member of Group 1, an excellent violin student, produced the violin notes in the earlier tests and at first during the practice periods. Later the writer was the experimenter and the observer practiced herself on the violin.

*Test IV* During the middle of the year all the members of Group 1 were similarly tested once on the pipe-organ. Four stops were used: the flute, the diapason and the viola stops with 60 notes each, and the oboe stop with 42 notes.

*Test V* At the close of the practice series with Group 1, nine of these observers attempted to reproduce the pitch of  $c^1$  and  $a^1$  singing into a Seashore tonoscope. Three trials were made for each note at wide intervals apart. The subject was instructed not to 'feel around' audibly with her voice to reach the pitch, as for example, to sing the extremes of her vocal range and estimate the pitch by that means. She was to sing directly with but slight modifications the pitch which she believed to be approximately  $c^1$  or  $a^1$ . All of the subjects had had opportunity during the year to use the tonoscope, which had been adapted by Professor D. C. Rogers for the high pitches of women's voices, and gave readily clear readings for these pitches.

*Test VI (a)* This test extended through the period following the spring vacation to the end of the college year. The eighty members of Group 2 were observers in this practice series. It was necessary that the procedure differ somewhat with so large a number of observers, hence the experiment was adapted for a group test in which about twenty observers could take part at one time. They sat with their backs to the piano facing the card keyboard and kept themselves a written record of their 88 responses. They were asked to attempt to place the note as near as possible to its actual position in every case.

The experimenter (the writer) presented the notes at fixed intervals, one every twenty seconds,<sup>3</sup> sounding each several times during the first 10 seconds, then pausing 10 seconds before striking the note. This made a trial last exactly 28 minutes. At the end the observers wrote introspections re-

---

<sup>3</sup>This interval seemed to be a fair average judging from the results of Group 1, which was allowed unlimited time.

garding the various mental processes involved in the judgments. These were read each week and criticized from the point of view of accurate psychological analysis.

The observers were given two preliminary trials without practice and six other trials at intervals of a week apart, following two ten minute periods of practice. The same Knabe piano was used both for the weekly practice and the group sittings. The correct notes, the semitone errors and the octave errors were designated in red after each trial to serve as an index to the observer of her achievement. Suggestions as to advisable methods of practice, based on results from Group 1, were made to this group. Further each observer was assigned a certain note<sup>4</sup> with which she was to make special efforts to become familiar. It was recommended that she think of her own note in a definite setting as belonging to a certain chord or scale, or possibly as the beginning note in some piece. They were to acquaint themselves with the general range of the piano and the tone-quality of each octave, and endeavor in some way to fix the individual notes in their memories, so that taking notes at random they could assign to them names with some degree of accuracy.

At the close of the practice series the following questionnaire was given to the observers from Groups 1 and 2:

### MEMORY FOR PITCH.

#### Training.

If you were to be graded on a scale of ten, where would you place yourself for (1) musical ability, (2) musical appreciation?

What instruments have you studied, and how long?

What vocal training have you had?

Have you taken ear training or harmony courses?

Have you a good, fair or poor memory for music?

What was your threshold for pitch discrimination?<sup>5</sup>

Are other members of your family musical or unmusical?

Type of Imagery.<sup>6</sup>

<sup>4</sup>The observers' names were arranged alphabetically according to class, and the notes were distributed to them in order from A<sup>2</sup> to f<sup>♯4</sup>.

<sup>5</sup>The pitch discrimination of a large number of the observers had been obtained at a class test earlier in the year.

<sup>6</sup>Adapted from Seashore's tests for imagery, *Elementary Experiments in Psychology*, 1908, pp. 107-110.

Answer the following questions by writing after the question the number which denotes the degree of vividness characteristic of your image. Follow the order: I 1, II 1, III 1, I 2, II 2, III 2, etc. Compute the averages for each type of imagery.

Degrees of Vividness.

0. No image at all.
1. Very faint.
2. Faint.
3. Fairly vivid.
4. Vivid.
5. Very vivid.
6. As vivid as in perception.

I. Visual.

1. Can you imagine the color of (a) a red rose? (b) a green leaf? (c) a yellow ribbon? (d) a blue sky?
2. Can you imagine the brightness of (a) a white teacup? (b) a black crow? (c) a gray stone? (d) the blade of a knife?
3. Can you image the form of (a) the rose? (b) the leaf? (c) the teacup? (d) the knife?
4. Can you form a visual image of (a) a moving express train? (b) your sharpening of a pencil? (c) an up and down movement of your tongue?
5. Can you image simultaneously (a) a group of colors in a bunch of sweet peas? (b) Colors, forms, brightnesses and movements in a landscape view?
6. Can you compare in a visual image (a) the color of cream and the color of milk? (b) the tint of one of your finger-nails with that of the palm of your hand?
7. Can you hold fairly constant for ten seconds (a) the color of the rose? (b) the form of the rose?

II. Auditory.

1. Can you image the sound of (a) the report of a gun? (b) the clinking of glasses? (c) the ringing of church bells? (d) the hum of bees?
2. Can you image the characteristic tone-quality of (a) a violin? (b) a cello? (c) a flute? (d) a cornet? (e) an organ? (f) an orchestra?
3. Can you repeat in auditory imagery the air of (a) the Star Spangled Banner? (b) the Marseillaise?



4. Can you form auditory images of the intensity of a violin-tone (a) very strong? (b) strong? (c) weak? (d) very weak?
5. Can you form auditory imagery of the rhythm (a) of men's marching to the beat of a drum? (b) "Dixie" or other air played? (c) "Tell me not in mournful numbers" spoken by yourself?
6. Can you form auditory images of (a) low tones on the piano, high tones, those in the middle range? (b) your own note on the piano? (c) middle c?

### III. Motor.

1. Can you imagine, in motor terms, yourself (a) rocking in a chair? (b) Walking down a stairway? (c) Biting a lump of sugar? (d) Clenching your fist? (e) Dancing?
2. Does motor imagery arise in your mind when you recall (a) cutting cloth with a large pair of scissors? (b) a facial expression of fear? (c) two boys on a teeter-board?
3. Aside from the actual inceptive movements, do you get motor imagery when recalling (a) a very high tone? (b) a very low tone? (c) middle c? (d) words like "Paderewski," "bubble," "tête à tête?"
4. Can you form a motor image of (a) the weight of a pound of butter? (b) your speed in running a race? (c) the motion of a boat? (d) being swung around rapidly in a chair or in a swing?

### Method.

On hearing the note did you try to recognize the octave first, then the note, or did you try to name the note first, then place it in an octave, or were the two simultaneous? In all cases?

Did the note to be judged seem to have a setting in an imagined chord, scale or keyboard? If so, did you think of the chord, scale or keyboard in visual, auditory, kin-aesthetic or any other terms?

Did recognition ever come instantly, without intervening images or associations?

To what extent were you conscious of overtones?

Did you notice perseveration of preceding tones? Did



you judge the pitch of one tone in reference to the pitch of the preceding tone?

Did you notice a difference in quality between the sound of a note immediately after it had been struck and after it had sounded a while? Did you prefer to judge at the initial sound of the note or after it had been sounded a few moments?

Explain fully what meaning  $c$  has for you. Do you think of it as having distinctive quality, coloring, setting, emotional value, or relation to other notes? Go through each of the twelve notes in this way: ( $c$ ,  $c\sharp$ ,  $d$ ,  $d\sharp$ ,  $e$ ,  $f$ ,  $f\sharp$ ,  $g$ ,  $g\sharp$ ,  $a$ ,  $a\sharp$ ,  $b$ ). How is  $c^2$  related introspectively to  $c^1$ ?

What sort of associative devices did you attempt?

Name some of the interferences which prevented the formation of these new associations, and some of the facilitations which aided their formation.

Did you feel special confidence in judging any tone or tones? Which ones? Did your practice aid in this?

By what method do you think you reached the best results?

What part of the keyboard did you find hardest? Which easiest? Why?

How would you criticise the length of interval or the manner of striking the notes in your own case?

Did you find the first, second or last third of the judgments in one sitting easier or harder than the others? Why?

Did you find the experiment interesting, indifferent or uninteresting?

(b) One year after the close of the practice series in which Groups 1 and 2 took part, twenty observers from Group 2 selected on the basis of availability and two observers from Group 1 were given similar tests as before. The members of Group 2 were given the group test and the members of Group 1 were tested individually with unlimited time for response. Introspections regarding the judging process were again written after the identifications were made.

(c) This same year Group 3, consisting of three especially selected music students, undertook a practice series similar to that of Group 2, in which the time for presentation and response was limited to twenty seconds. A concert grand piano located in the office of a member of the musical depart-

ment was used for these trials. Introspections were recorded at the end of each sitting and the questionnaire was filled out at the close of the practice series.

*Test VII* The ability to recognize the fundamentals of tonic chords evenly struck in the twelve different keys between  $c^1$  and  $c^8$  and the improvement in this ability by practice was investigated during the summer of 1918.<sup>7</sup> Two observers (a new observer, B, who possesses marked musical ability and observer G, who had taken part in the tests of Group 1) carried this on as a practice experiment. Each of the twelve chords was presented twice at a sitting, all twenty-four being arranged at relatively wide tonal intervals apart and given after pauses of a minute, in order to eliminate as far as possible relative judgments. During the following winter one of the music students from Group 3 did the experiment as a single test in identification following the same procedure as was used for the previous series. Introspections were recorded during the practice tests.

*Test VIII* Since the psychological laboratory at Smith College had at this time ready for use a set of instruments for detecting slight movements during the thinking process<sup>8</sup> it seemed worth while to apply these instruments to the field of pitch judgments. Consequently while two observers, members of Group 1, judged the pitch of 31 violin notes selected at random and played at uniform intervals of fifteen seconds, kymograph records were made of thoracic breathing, horizontal movements of the throat, tongue movements and the movements and changes in volume of the left arm.<sup>9</sup> Both observers were familiar with the quality of sound and the technique of the violin and with the rather elaborate apparatus necessary for taking these records.

*Test IX* Professor Seashore suggested to the writer that an accurate test for detecting genuine absolute pitch memory would consist in presenting the observer with a tuning fork early in the morning before he had heard any music which might set up a temporary relative "set" and ask him to judge

---

<sup>7</sup>For these experiments a relatively new baby grand piano was used.

<sup>8</sup>See R. S. Clark, An Experimental Study in Silent Thinking, *Archives of Psychology*, No. 48, 1922.

<sup>9</sup>About two yards of smoked paper was stretched from a roller over a very slowly rotating kymograph drum, so that it was possible to proceed without break through the test which lasted nearly eight minutes.

its pitch. Smaller differences than musical semitone intervals ought to be judged accurately if the observer's memory is strictly a memory for absolute pitch. A set of tuning forks containing the standard fork of 435 v. d. and the ten increment forks 0.5, 1, 2, 3, 5, 8, 12, 17, 23 and 30 vibrations higher than the standard furnished the material for the test. Each morning for 66 fairly consecutive mornings, as far as was possible in suburban surroundings before hearing any musical sounds, observer G from Group 1 was given one of the eleven tuning forks, to sound as often and as long as was desired, and to pass judgment in regard to which of the eleven forks it might be. Each fork was presented six times in random order during the experiment. The observer was already familiar with the sound and range of the tuning forks and after each daily identification sounded all the various forks and listened carefully to their tones. The results of the test were not made known to the observer until the close of the experiment.



## IV

### DATA AND DISCUSSION

Since many of the investigations of this kind had begun their study of memory for tones by using first a small number of notes and gradually increasing the number, it seemed advisable to include this method as well as one by which the entire keyboard is attempted at each trial. After continuing the two methods in Group 1 for several months the piecemeal method was abandoned for these reasons. It was less interesting to the observers and called forth less responsiveness from them. The narrow range allowed little opportunity for variety in forming chord combinations in practice and for testing octave similarities and differences. This method also complicated the procedure of the tests somewhat. The different notes could not be presented at relatively short time intervals without distractions, for relative pitch operated too readily within the narrow limits. It made the interpretation of results a more complex affair which was disadvantageous in an experiment with a large group. Therefore when weekly tests on eighty observers were undertaken, simplicity in procedure became essential.

1. *Tables, graphs and correlations for the experiments using the 88 notes of the piano.*

TABLE I<sup>1</sup>

Showing the Frequency-Distribution of Errors measured in Semitones and averaged from the Total Number of Trials of Groups 1, 2 and 3 (92 Observers).

Average Error Per Note in Semitones	Number of Observers	Average Error Per Note in Semitones	Number of Observers
0	0	3.97	20
0.56	0	4.54	18
1.13	1	5.11	22
1.70	1	5.68	11
2.27	1	6.25	6
2.84	4	6.81	0
3.40	7	7.38	1

a. *Average error frequency* The results given in Table I show that the errors of the observers from Groups 1, 2 and 3 approximate a normal curve of distribution. (See Figure 1, page 57) No sharp break occurs between any two divisions although the observers range from those who claim to be tone-

<sup>1</sup>Tables I through XII and the corresponding Figures 1 through 8 are based on the results from the tests using the 88 notes of the piano for identification.



deaf and those who had practically no musical education, through all stages including the unmusical who have received good training and the musically inclined who have had little training, to the especially promising students of music. The average error for the various members of these groups is less than an interval of a perfect fourth per note, for the poorest observer is a little over a minor sixth and for the best observer slightly over a semitone. Evidently the ability to identify notes is not an all-or-none trait, but may be found in all grades of accuracy. Even the least successful observers are able to assign positions to notes within certain relatively narrow limits.

TABLE II

Showing the Frequency-Distribution of Errors, measured in Semitones, from the First Trials of Groups 1, 2 and 3. (88 observers)

Average Error Per Note in Semitones	Number of Observers	Average Error Per Note in Semitones	Number of Observers
0.56	0	4.54	8
1.13	0	5.11	22
1.70	0	5.68	18
2.27	2	6.25	12
2.84	0	6.81	7
3.40	5	7.38	4
3.97	7	7.95	3

b. *Error frequencies for the first and the last trials* The curve showing the frequency of errors for the first trial (See Figure 2, page 58) assumes also the general form of the probability curve. After a period of practice the distribution remains approximately the same, although the central point has moved about 100 semitones nearer the line of no errors. (See also Tables II and III)

The improvement per person, however, was not as uniform as this graph might indicate, for using the Pearson formula

TABLE III

Showing the Frequency-Distribution of Errors, measured in Semitones, from the last Trial of the Practice Series of Groups 1, 2 and 3. (88 observers)

Average Error Per Note in Semitones	Number of Observers	Average Error Per Note in Semitones	Number of Observers
0	1	3.97	21
0.56	0	4.54	12
1.13	1	5.11	12
1.70	2	5.68	9
2.27	6	6.25	2
2.84	10	6.81	1
3.40	11	7.38	0

for the coefficient of correlation<sup>2</sup> to obtain the degree of relationship between initial and final trials in the practice series as measured by the size of the error, it is found that in Groups 2 and 3,  $r = +.439$  p. e. .06, and in Group 1,  $r = -.132$  p. e. .219. Certain individuals improved rapidly after a poor beginning, others who did better in their first performance gained slowly in comparison. This is most marked in Group 1. An important reason for the difference in the correlation of Groups 2 and 3 and that of Group 1 lies in the fact that the period of practice varied greatly among the observers in this latter group, some continuing this special form of the experiment for only six trials, others for fifteen to twenty-four trials. The longer the periods of practice the more chance for accentuating the factors of individual difference. Also those who began at a high level of ability had less chance for conspicuous improvement than the less able ones.

c. *Average frequency for correct notes* Greater deviation from the normal curve of distribution is shown in Table IV, Figure 3, based on individual records for correct notes, than in the curve for error distribution. The curve is skewed toward the side of little ability and is very irregular at the opposite side. On the whole the observers in Group 1 whose practice extended over longer periods average much better than the observers in Group 2 whose practice lasted six weeks. The relationship between the average number of correct notes and accuracy as measured by the size of the average error proved to be  $r = +.602$  p.e. .047 for Groups 2 and 3, and  $r = +.895$  p. e. .044 for Group 1. It will be remembered that the tests as given to Groups 2 and 3 were in the form of group tests with a definite time limit for response. The observers in Group 1, on the contrary, were tested individually and given unlimited time for response. This factor together with the variation in length of the practice series would tend to produce a greater degree of correspondence between the general correctness of response as measured negatively by the size of the

<sup>2</sup>The following formulae for the Pearson coefficient of correlation and the probable error were used in all the correlations which follow:

$$r = \frac{\sum (x^1 \cdot y^1) - n \cdot \text{Error}_x \cdot \text{Error}_y}{n \left( \frac{\sum (x^1)^2}{n} - \text{Error}_x^2 \right)^{\frac{1}{2}} \left( \frac{\sum (y^1)^2}{n} - \text{Error}_y^2 \right)^{\frac{1}{2}}}$$

$$\text{P.E.} = .6745 \frac{1 - r^2}{n^{\frac{1}{2}}}$$

TABLE IV

Showing the Frequency-Distribution of Correct Notes averaged from the Total Number of Trials of Groups 1, 2 and 3. (90 observers)

Average Number of Correct Notes	Number of Observers		
	Group 1 (Unlimited Time)	Group 2 (Limited Time)	Group 3 (Limited Time)
3		1	
4		1	
5		10	
6		17	
7		16	
8		10	
9		9	
10	1	6	
11	4	3	
12		2	
13		1	
14	1		
15	1	1	1
17	1		
21			1
23		1	
40	1		
72			1
Total	9	78	3

errors and positively by the number of correct notes. Since this high correlation is found between the two methods of grading, it makes it possible to use either of the two more or less interchangeably as a basis for estimating an individual's record.

d. *Frequencies of correct notes in the first and the last trials* The Frequency-Distribution of correct notes in the first trial is also skewed toward the poorer extreme. (See Figure 4, Tables V and VI.) Practice draws the entire curve over a little to the opposite side but does not change its original form. It seems even to accentuate individual differences by increasing the number of fluctuations in the curve for the last trial. This is in contrast to the fairly normal curve of distribution found in the curves of average error. This greater variety on the side of the curve representing greater achievement may perhaps be explained as the result of the advantages in general education and musical training which the observers in these tests have had.



TABLE V

Showing the Distribution of Correct Notes in the First Trial of the Practice Series of Groups 1, 2 and 3. (90 observers)

Number of Correct Notes in First Trial	Number or Observers	Number of Correct Notes in First Trial	Number or Observers
1	1	10	3
2	3	11	6
3	7	12	3
4	9	13	2
5	14	14	1
6	13	16	1
7	9	18	1
8	13	63	1
9	3		

Although the members of Groups 1 and 2 represent as wide a sampling as is possible in a college community, it is however true that that community is already narrowly selective.

Improvement in naming correctly the individual notes is steady in the group as a whole and alters but little the form of the curve, but individual improvement is at times somewhat uneven. The correlation between the number of correct notes per individual in the first and in the last trials for Groups 2

TABLE VI

Showing the Distribution of Correct Notes in the Last Trial of the Practice Series of Groups 1, 2 and 3. (90 observers)

Number of Correct Notes in Last Trial	Number of Observers	Number of Correct Notes in Last Trial	Number of Observers	Number of Correct Notes in Last Trial	Number of Observers
3	1	11	8	20	1
4	2	12	5	21	1
5	6	13	2	22	1
6	7	14	2	24	1
7	12	15	2	25	1
8	12	16	4	36	1
9	11	18	2	44	1
10	5	19	1	83	1

and 3 is:  $r = +.845$  p. e. .031 and for Group 1:  $r = +.516$  p. e. .266. On the whole those who begin well are apt to be the ones to finish well, especially when the practice periods are of equal length. Where the length of the practice series varies greatly as occurred in Group 1, the record for the first trial cannot be taken as an accurate gauge of the results of different individuals after unequal intervals of practice.

e. *Frequency of interval errors* It was of great interest to find the relative frequency of the different interval errors taken from over 700 tests in which the 88 notes of the piano



were used. (See Figure 5, Table VIII) Underestimations of pitch are fairly consistently less than overestimations. The largest number of misjudgments were only a semitone too high or too low. The next lower frequency was found for the whole

TABLE VII

Showing the Total Frequency of Interval Errors up to Two Octaves obtained from all Records (89 Observers).

Size of + Interval Errors	Total Frequency	Size of — Interval Errors	Total Frequency
No Error	7621		
Minor 2nd	5821	Minor 2nd	4462
Major 2nd	4354	Major 2nd	3488
Minor 3rd	3357	Minor 3rd	2869
Major 3rd	2866	Major 3rd	2337
Perfect 4th	2564	Perfect 4th	1869
Augmented 4th	2063	Augmented 4th	1595
Perfect 5th	1824	Perfect 5th	1331
Minor 6th	1470	Minor 6th	1094
Major 6th	1305	Major 6th	904
Minor 7th	1078	Minor 7th	797
Major 7th	983	Major 7th	763
Perfect 8th	917	Perfect 8th	692
Minor 9th	793	Minor 9th	368
Major 9th	489	Major 9th	285
Minor 10th	345	Minor 10th	213
Major 10th	283	Major 10th	177
Perfect 11th	204	Perfect 11th	113
Augmented 11th	123	Augmented 11th	90
Perfect 12th	110	Perfect 12th	67
Minor 13th	68	Minor 13th	71
Major 13th	54	Major 13th	38
Minor 14th	38	Minor 14th	29
Major 14th	35	Major 14th	30
2 Octaves	24	2 Octaves	13

tone, then the minor third, the major third and so on down step by step. The larger the interval error the less frequently it occurred. Instead of a distinct rise in frequency for octave errors the curve continues downward, a little more slowly at that point than a few intervals before or a few intervals afterward. In the records of 31 observers out of 89, the frequency for the positive octave error rises one point or more above the frequency for the error of the major 7th interval. In only nine cases is this rise conspicuous and in only three cases does it occur among the negative interval errors. In four cases the frequency for the octave error rises above all preceding frequencies except, in one case, that of the semitone error, and, in the other three cases, that of the whole tone error. Not the slightest evidence is found in any record of a preference for errors of a fourth or a fifth, as is sometimes claimed to be the

case. The major 3rd has the modal frequency of error in two cases, the minor 3rd in three cases, the major second in seventeen cases, and the minor second in all other cases (numbering 67).

f. *Average time per note* The time for each response of the observers in Group 1 was taken by means of a stop watch. The average time per note in each trial was found for the five observers who continued the experiment by the same method throughout the academic year, and recorded in Table VIII. (See Figure 6, page 62.) Wide variations from trial to trial are shown in each record. C in her introspections admits that she followed very different methods on different occasions. This would account for fluctuating from 13.6 seconds per note to 18.7 seconds to 10.4 seconds. The desire to improve upon the accuracy of past records had a tendency to lengthen the time, for instead of making an immediate quick judgment the observers weighed their decisions several moments to assure themselves of the approximate correctness. Fatigue and changes in the level of the attention affected the speed of response. Many of the tests were given late in the afternoon when concentration on a task which taxed one's auditory perceptions to so unusual a degree was difficult.

TABLE VIII

Showing the average Time per Note in each Trial of the Practice Series. (5 Observers from Group 1)

Trials in Practice	Series	Average Time per note in seconds				
		B	C	E	F	G
1		11.5	4.4	6.0	10.9	4.9
2		14.2	7.4	6.2	10.9	4.9
3		13.7	9.5	6.6	12.4	7.4
4		13.5	10.8	6.3	14.0	7.2
5		17.9	10.9	6.5	18.9	9.1
6		21.1	14.9	5.6	22.3	13.7
7		9.9	17.3	8.2	19.8	7.9
8		11.7	17.6	8.4	17.2	6.0
9		15.9	13.6	9.3	20.5	7.1
10		11.6	18.7	8.2	16.5	6.5
11		12.8	10.4	8.8	19.0	5.1
12		14.9	12.8	7.7	20.6	5.8
13		9.9	14.2	8.9	20.8	4.7
14		13.1	14.8	8.5	19.2	5.8
15		13.8	13.2	9.0	16.3	4.5
16		16.7	11.6	7.6		5.6
17		23.6	12.6	8.3		5.9
18		21.2	11.6			3.8
19		20.7				5.0
20						5.9
21						5.9
22						6.5

The observer with the best record of accuracy in the test had the lowest average time, but the observer with the next best record had the highest average time. The observer who was second best in time had the least successful record of accuracy of the five observers. (See Table IX.) Comparing the average time and average accuracy measured by the amount of error in each trial of the practice series of these five observers, the coefficient of correlation in each case is as follows:

$$\begin{aligned} B\ r &= +.028\ \text{p.e.}\ .153 \\ C\ r &= +.528\ \text{p.e.}\ .152 \\ E\ r &= +.807\ \text{p.e.}\ .056 \\ F\ r &= +.308\ \text{p.e.}\ .014 \\ G\ r &= -.216\ \text{p.e.}\ .136 \end{aligned}$$

In the case of E, the longer the time taken for response the fewer the errors. There is but slight indication of this for C and F, none at all for B and the barest indication that the reverse is true in the case of G. On the whole the time per response is a very inadequate gauge of ability in these tests. Correlating the time per individual response with the degree of accuracy in G's trial 17,  $r = -.037\ \text{p.e.}\ .071$ . In this trial 50 of the 88 notes were correct and only one octave error occurred which might over-emphasize the size of individual errors.

*g. Average time for correct and incorrect notes* The average times for correct and for incorrect notes found from all the trials of the five observers show a slightly longer time for the incorrect responses in four instances and a longer time for the correct responses in one instance. (See Figure 7, page

TABLE IX

Showing the Average Time for all Notes, for Correct Notes and for Incorrect Notes, the Size of the Average Error and the Average Number of Correct Notes. (5 Observers)

Observer	Average Time Per Note in Seconds	Average Time For Correct Notes in Seconds	Average Time For Incorrect Notes in Seconds	Size of Average Error per Note Measured in Semitones	Average Number of Correct Notes
B	15.1	14.1	15.4	4.1	11.2
C	12.6	12.0	12.7	3.4	15.3
E	7.7	6.9	7.8	3.4	10.5
F	17.3	17.6	17.3	2.0	17.5
G	6.3	5.4	7.2	1.7	40.6



63.) This last observer admits that she made use of relative pitch in a large measure; and in consequence of her round-about method all her times are long. It is natural that the correct judgments should be the somewhat quicker responses. The observers knew their own limitations fairly well and would pause to consider certain notes in the keyboard which were invariably difficult to judge and try to find in them distinguishable pitch or other tonal characteristics which might facilitate judgment on later occasions.

*h. Individual practice curves* The reduction in error in the individual practice curves of the five observers of Group 1 is exceedingly irregular. (See Table X, Fig. 8.) The interruptions caused by the vacation and mid-year examination periods are responsible in part for this, although variations in method and attitude are more important factors. All observers made definite initial progress. With possibly one exception no definite plateaus occur at any place, and the curves indicate on the whole a tendency toward improvement which at the end of the test had by no means reached a level. F alone did not progress in her last five trials. Her method, which was based mostly on the use of relative pitch whenever possible,

TABLE X

Showing the Amount of Error per Note in each Trial of the Practice Series of Five Observers from Group 1.

Trials	Average Error per Note measured in Semitones				
	B	C	E	F	G
1	5.6	4.8	5.2	3.9	2.5
2	4.4	4.8	4.3	3.7	2.8
3	5.6	5.4	4.3	3.8	2.1
4	5.8	3.8	4.6	3.2	2.0
5	5.4	4.4	3.6	3.1	1.9
6	4.9	3.1	3.8	3.4	2.4
7	5.2	4.7	3.6	2.5	2.1
8	4.1	3.4	2.6	2.4	1.4
9	4.6	3.0	3.4	3.1	1.9
10	3.1	2.3	3.0	2.3	1.3
11	2.8	2.7	2.6	3.4	1.2
12	3.6	3.4	3.0	2.8	1.8
13	3.6	3.7	3.1	2.5	2.4
14	3.7	2.2	2.8	2.4	2.5
15	3.8	2.9	3.4	3.2	1.4
16	4.0	2.6	3.2		.7
17	3.5	2.2	2.2		1.1
18	2.9	2.5			1.6
19	2.1	1 Year Later	3.4		Continued during the Summer
20					1.9
21					1.0
22					1.4
23					1 Year Later



or what seemed to be non-essential associative devices, may have hindered her improvement. G's lack of progress during the summer is partly the result of the use of a piano with a wholly different quality of sound, and partly the result of the inability to sustain a high level of attention on oppressively hot days.

*2. Additional Correlations* In order to study the effect of the amount of time spent in practice on improvement, and the effect of previous musical training on improvement, the Pearson coefficient was obtained from the data given by 84 members of the three groups. Improvement was measured in an arbitrary manner. Because of the fact that those who started out poorly had much more opportunity to improve than those who began well, a definite advantage increasing in arithmetic progression was added to the actual amount of improvement of the last trial over the first trial in each case.<sup>3</sup> The correlation between this arbitrary amount of improvement and the amount of time spent in practice gave  $r = +.391$  p.e. .061; the correlation between improvement and previous musical training gave  $r = +.205$  p.e. .049. In both cases the positive relationship is barely existent. Evidently frequency of hearing the piano notes is alone no guarantee that memory for absolute tone will result. It must be noted that the specifically directed practice for this experiment gave a slightly higher correlation with improvement than the general musical training of past years gave.

The average degree of accuracy per person was correlated with the number of years of musical training and gave  $r = +.436$  p.e. .055. This shows some tendency for the musically trained to make a smaller amount of error on the whole than the musically untrained.

Sixty of the observers had taken part in a group test for pitch discrimination, in which the standard set of tuning forks was used. The standard fork and each of the ten increment forks was repeated ten times and 70% accuracy was taken as the threshold for discrimination. The results from

---

<sup>3</sup>The largest total amount of error found in trial 1 was 745 semitones. Using this as a starting point, those having an error of 745 and 744 in their first trials had zero added to their amount of improvement. The observers starting with 743 and 742 errors measured in semitones received an increment of one point each, those starting with 741 and 740 errors an increment of two points each, and so on with regular increase to the observer having the smallest amount of error which proved to be 223. To the improvement in this case was added 261 points.

these tests were correlated with the average error of these sixty observers and gave  $r = +.376$  p.e. .073. That means roughly that a low threshold has a slightly better chance for accompanying a record with few errors than a higher threshold. The best records were made for the most part by those with the keenest ears.

### 3. *Additional tables*

a. *Average error in the piano test one year later* One year after the close of the tests with Groups 1 and 2, sixteen members of the groups were retested in a similar manner. Observers from Group 2 were given the group test with a time limit for response and those from Group 1 were tested individually with unlimited time.<sup>4</sup> An examination of the results given in Table XI shows that one-fourth of the observers did better as measured by average error the following year than at the last trial of the previous year. One-half made a better record in the retesting than the average of their practice series records, and fourteen of the sixteen in their retesting improved on their records for the first trials of the series. The correlation between the size of the error of the last trial of the practice series and the size of the error at the retesting gives  $r = +.68$  p.e. .09. This indicates that the amount retained

TABLE XI

Showing the Size of the Error measured in Semitones of a Trial One Year after the Close of the Practice Series, Retesting the Ability to Identify the 88 Piano Notes, and its Relation to Previous Errors.

Observers	Error Per Note	Error Per Note in First Trial	Error Per Note in Last Trial	Error Per Note One Year Later
Group 1 G	1.7	2.5	1.4	1.5
C	3.4	4.8	2.5	3.3
Group 2 Rn	3.6	6.2	2.5	3.1
M	3.9	7.7	3.3	3.2
Bt	4.0	4.8	2.8	4.2
K	4.0	5.1	3.6	3.4
Mc	4.2	5.0	3.3	3.9
T	4.3	5.5	2.8	5.1
Da	4.3	5.3	3.4	4.4
Hs	4.4	5.8	3.2	4.6
Ce	4.5	6.3	4.5	5.1
Hk	4.8	5.2	4.3	3.9
S	5.2	7.6	2.6	5.3
Kp	5.2	6.1	5.5	6.6
Hn	5.4	5.1	5.9	5.9
P	5.7	5.7	5.0	5.4

<sup>4</sup>These observers were selected simply on the basis of availability, consequently include good, fair and poor observers.

over a period without practice is to some degree proportional to the degree of ability acquired from the practice. Four of the six observers who said they had not attempted to identify notes during the year just past, made records poorer than the results averaged from the total number of trials of the practice series. Six of the ten observers who had made some attempts at tonal recognition in the year's interval made better records than the average record of their practice series.

b. *Individual frequency of correctness for the 88 notes on the piano* The frequency with which the correct names were given to each note on the piano keyboard is given in Table XII. The *c*'s have the highest frequency of correctness, then in order come *a*, *b*, *g*, *d*, *f*, *e*, *a*<sup>♯</sup> (or *b*<sub>b</sub>), *g*<sup>♯</sup> (or *a*<sub>b</sub>) and *d*<sup>♯</sup> (or *e*<sub>b</sub>), *c*<sup>♯</sup> (or *d*<sub>b</sub>), and *f*<sup>♯</sup> (or *g*<sub>b</sub>). *C*<sup>5</sup> has the largest number of correct responses. The next eleven in order are: *A*<sub>2</sub>, *b*<sup>4</sup>, *a*<sup>4</sup>, *c*<sup>1</sup>, *c*<sup>2</sup>, *B*<sub>2</sub>, *a*<sup>♯4</sup> (or *b*<sub>b</sub><sup>4</sup>), *A*<sup>♯2</sup> (or *B*<sub>b</sub><sup>2</sup>), *g*<sup>4</sup>, *D*<sub>1</sub> and *d*<sup>1</sup>. The twelve in order with the smallest number of correct responses are: *c*<sup>♯3</sup> (or *d*<sub>b</sub><sup>3</sup>), *F*<sup>♯</sup> (or *G*<sub>b</sub>), *g*<sup>♯3</sup> (or *a*<sub>b</sub><sup>3</sup>), *G*<sup>♯</sup> (or *A*<sub>b</sub>), *D*<sup>♯</sup> (or *E*<sub>b</sub>), *d*<sup>♯3</sup> (or *e*<sub>b</sub><sup>3</sup>), *C*<sup>♯</sup> (or *D*<sub>b</sub>), *c*<sup>♯</sup> (or *d*<sub>b</sub>) and *d*<sup>♯</sup> (or *e*<sub>b</sub>), *f*<sup>♯1</sup> (or *g*<sub>b</sub><sup>1</sup>) and *c*<sup>♯4</sup> (or *d*<sub>b</sub><sup>4</sup>), and *f*<sup>♯3</sup> (or *g*<sub>b</sub><sup>3</sup>). Excluding the three notes of the subcontra octave and *c*<sup>5</sup>, the seven octaves arrange themselves according to frequency of correct responses as follows: four-accented octave, once-accented (or middle) octave, contra octave, twice-accented octave, small octave thrice-accented, and great octave. It is to be expected that the results from these groups of observers who did not at the start possess memory for absolute tone will be markedly different from results obtained from the selected, especially gifted group. However, if we leave out of consideration the subcontra octave and the four-times accented octave (which by pure chance the observers might name correctly more frequently than the central octaves), and give the remaining five their rank in this respect, the results from these groups coincide exactly with Baird's results (see page 21), namely:

C-B	c-b	c <sup>1</sup> -b <sup>1</sup>	c <sup>3</sup> -b <sup>3</sup>	c <sup>4</sup> -b <sup>4</sup>
5	3	1	2	4



TABLE XII

Showing Individual Relative Frequency of Correctness for the 88 Notes on the Piano (90 Observers).

Sub- Contra Octave	No. Cor- rect	Contra Oc- tave	No. Cor- rect	Great Oc- tave	No. Small Cor-Octave	No. Cor.	Once- Acc. Octave	No. Twice- Acc. Octave	No. Three- Acc. Octave	No. Four- Cor- rect	No. Five- Acc. Octave	No. Freq. Note irre. of Octave
		C <sub>1</sub>	72	C	53	51	c <sup>1</sup>	c <sup>2</sup>	c <sup>3</sup>	47	c <sup>5</sup>	95.5
		C <sup>#</sup> <sub>1</sub> D <sub>b1</sub>	65	C <sup>#</sup> D <sub>b</sub>	25	27	c <sup>#</sup> db <sup>1</sup>	c <sup>#</sup> db <sup>2</sup>	c <sup>#</sup> db <sup>3</sup>	16	c <sup>#</sup> db <sup>4</sup>	35.0
		D <sub>1</sub>	77	D	42	45	d <sup>1</sup>	d <sup>2</sup>	d <sup>3</sup>	66	d <sup>4</sup>	59.1
		D <sup>#</sup> <sub>1</sub> E <sub>b1</sub>	54	D <sup>#</sup> E <sub>b</sub>	23	27	d <sup>#</sup> eb <sup>1</sup>	d <sup>#</sup> eb <sup>2</sup>	d <sup>#</sup> eb <sup>3</sup>	24	d <sup>#</sup> eb <sup>4</sup>	38.0
		E <sub>1</sub>	72	E	42	36	e <sup>1</sup>	e <sup>2</sup>	e <sup>3</sup>	49	e <sup>4</sup>	54.6
		E <sup>#</sup> <sub>1</sub>	65	F	47	54	f <sup>1</sup>	f <sup>2</sup>	f <sup>3</sup>	56	f <sup>4</sup>	57.4
		F <sup>#</sup> <sub>1</sub> G <sub>b1</sub>	36	F <sup>#</sup> G <sub>b</sub>	18	33	f <sup>#</sup> gb <sup>1</sup>	f <sup>#</sup> gb <sup>2</sup>	f <sup>#</sup> gb <sup>3</sup>	30	f <sup>#</sup> gb <sup>4</sup>	32.3
		G <sub>1</sub>	63	G	41	74	g <sup>1</sup>	g <sup>2</sup>	g <sup>3</sup>	82	g <sup>4</sup>	61.6
		G <sup>#</sup> <sub>1</sub> A <sub>b1</sub>	32	G <sup>#</sup> A <sub>b</sub>	21	58	g <sup>#</sup> ab <sup>1</sup>	g <sup>#</sup> ab <sup>2</sup>	g <sup>#</sup> ab <sup>3</sup>	19	g <sup>#</sup> ab <sup>4</sup>	38.0
A <sub>2</sub>	175	A <sub>1</sub>	41	A	32	62	a <sup>1</sup>	a <sup>2</sup>	a <sup>3</sup>	49	a <sup>4</sup>	73.9
A <sup>#</sup> <sub>2</sub> B <sub>b2</sub>	86	A <sup>#</sup> <sub>1</sub> B <sub>b1</sub>	34	A <sup>#</sup> B <sub>b</sub>	33	50	a <sup>#</sup> bb <sup>1</sup>	a <sup>#</sup> bb <sup>2</sup>	a <sup>#</sup> bb <sup>3</sup>	36	a <sup>#</sup> bb <sup>4</sup>	50.1
B <sub>2</sub>	96	B <sub>1</sub>	32	B	38	48	b <sup>1</sup>	b <sup>2</sup>	b <sup>3</sup>	39	b <sup>4</sup>	62.8
Average	119		53.6		34.6	47.1	58.8	50.6		40.0	69.2	287



The assignment of individual notes to the observers in Group 2 had definite value. First they aided in making the task more concrete and stimulating. Little stress was laid on the notes after they were assigned, for it was thought that emphasis on individual notes might divert the attention from a study of the entire range. The observers, however, took readily to the idea of mastering a single note and frequently practiced the other notes in relation to this one. Second the results from the notes assigned illustrate the value of specially directed practice. In comparing the frequency of right responses for individual notes with that for the average note, it is striking that the assigned notes were correctly named approximately two and a half times as often as the average note.

TABLE XIII

Showing the Average Error per Note in Semitones, the Range of Positive and Negative Errors and the Median Position for each 3-Semitone Division of the 87 Piano Notes, based on 725 Trials from Groups 1, 2 and 3, and measured in Semitones.\*

3-Semitone Divisions	Average Error per Note in Semitones	Range of Errors in Semitones		Median Position per Note in Semitones
		Positive	Negative	
1 A <sub>2</sub> —	3.4	25	— 2	+1.14
2 C <sub>1</sub> —	3.8	26	— 5	+ .20
3 D <sub>1</sub> <sup>#</sup> —	3.6	27	— 8	— .89
4 F <sub>1</sub> <sup>#</sup> —	4.6	27	—11	+ .75
5 A <sub>1</sub> —	5.0	26	—14	+ .11
6 C —	5.1	26	—17	— .15
7 D <sub>1</sub> <sup>#</sup> —	5.1	26	—19	+ .72
8 F <sub>1</sub> <sup>#</sup> —	5.5	26	—22	+ .36
9 A —	5.4	27	—23	+ .41
10 c —	5.9	34	—25	+1.17
11 d <sub>1</sub> <sup>#</sup> —	5.3	30	—24	+1.22
12 f <sub>1</sub> <sup>#</sup> —	4.7	30	—23	+ .47
13 a —	4.0	22	—28	+ .13
14 c <sup>1</sup> —	3.8	24	—25	+ .04
15 d <sub>1</sub> <sup>#</sup> <sup>1</sup> —	4.0	29	—28	— .56
16 f <sub>1</sub> <sup>#</sup> <sup>1</sup> —	4.0	26	—28	— .55
17 a <sup>1</sup> —	4.1	27	—22	— .73
18 c <sup>2</sup> —	4.5	26	—25	+ .03
19 d <sub>1</sub> <sup>#</sup> <sup>2</sup> —	4.7	26	—26	+ .25
20 f <sub>1</sub> <sup>#</sup> <sup>2</sup> —	5.2	26	—27	+ .20
21 a <sup>2</sup> —	5.6	25	—34	+ .30
22 c <sup>3</sup> —	5.8	23	—26	+ .08
23 d <sub>1</sub> <sup>#</sup> <sup>3</sup> —	5.5	19	—27	+ .13
24 f <sub>1</sub> <sup>#</sup> <sup>3</sup> —	5.4	18	—33	— .64
25 a <sup>3</sup> —	5.3	14	—35	— .87
26 c <sup>4</sup> —	5.0	12	—35	+ .39
27 d <sub>1</sub> <sup>#</sup> <sup>4</sup> —	4.1	9	—27	+ .13
28 f <sub>1</sub> <sup>#</sup> <sup>4</sup> —	3.1	6	—25	+ .05
29 a <sup>4</sup> —	2.1	3	—23	— .74

\*c<sup>5</sup> was omitted in order to preserve the 3-semitone grouping.

c. *Average error per note, range of positive and negative errors and median position for each 3-semitone division of the 87 piano notes* The results from experiments in pitch discrimination show that absolute differences in discriminative ability measured in terms of vibration rates tend to remain constant over a limited range.<sup>6</sup> Hence an increase in keenness follows a rise in pitch when measured relatively, that is, in terms of a fractional part of a tone. These facts would lead one to expect a similar gain in accuracy in judgments of absolute pitch. The data, however, do not bear out this supposition. The correlation between average accuracy, based on the average error measured in semitones from 725 records, of the 35 'naturals' from the five octaves *C* to *c*<sup>4</sup> and the absolute size of the semitone interval measured in terms of vibration rates, gives  $r = -.24$  p.e. .024.<sup>7</sup> The accuracy remains fairly steady between *C* and *f*, where the average error is about five semitones, then improves between *g* and *e*<sup>2</sup>, *c*<sup>1</sup> having the smallest amount of error, namely, 3.6 semitones, and from *e*<sup>2</sup> decreases and remains with an error of slightly over five semitones up to *c*<sup>4</sup>. The median positions for the individual notes taken in groups of threes (Fig. 9) show remarkably little variation from their true positions.

Other factors were obviously operating against this admitted increase in discriminative ability. Frequency of hearing the notes in the middle range was undoubtedly a very important element. Loss of distinct changes in volume with rise in pitch might counteract the benefits derived from relative improvement in pitch discrimination. Rich tested for volume limens at three levels, 275 v.d., 550 v.d. and 1100 v.d., and from his results concludes that volume tends to follow Weber's Law.<sup>8</sup> Whatever the status of volume it proved influential in judgments of absolute pitch and a progressive lessening of its efficiency in these regions tested would affect the accuracy of pitch judgments.

The chief factor which can account for the sharp turns of the curve at *c* and at *c*<sup>3</sup> (Fig. 10) is the 'end-error.' The largest possible negative error at *c* is 27 semitones, and the largest

<sup>6</sup>C. E. Seashore, *The Psychology of Musical Talent*, 1919, p. 73.

<sup>7</sup>The overestimations and the underestimations per note were added together and correlated with the semitone interval above and below each respective note.

<sup>8</sup>G. T. Rich, *A Study of Tonal Attributes*. *American Journal of Psychology*, 1919, 50, pp. 149-153.

possible positive error at  $c^2$  is 24 semitones, owing to the extremes of the pianoforte. An error of approximately these amounts is found to occur fairly regularly throughout the intermediate range of notes. On approaching the extremes of the piano the size of the plus or minus error is necessarily reduced and the curve falls correspondingly.

d. *Average error and average time in the test with the pipe-organ* The experiments on the organ were undertaken with the idea of making a study of the effects of different timbres upon tonal judgment. Many difficulties were met in the attempt. The organ was less accessible than the piano and the range of the stops varied considerably, so that a comparison of results was difficult. Therefore these experiments were

TABLE XIV

Showing the Average Error per Note measured in Semitones and the Average Time per Note in Seconds for Three Observers in Judging the Notes from Organ Stops.

Observers	Organ Stops used	Average Error per Note Measured in Semitones	Average Time per Note in Seconds
C	Viola (60 notes)	3.2	17.3
	Diapason (60 notes)	3.5	9.7
	Oboe (40 notes)	3.4	13.4
	Flute (60 notes)	9.3	12.4
	Piano (Trial 1)	4.8	
	Experiment (Aver.)	3.3	12.6
Ht	Viola	3.7	13.6
	Diapason	4.4	7.3
	Oboe	6.4	7.1
	Flute	6.3	8.4
	Piano (Trial 1)	5.5	
	Experiment (Aver.)	3.6	10.2
B	Viola	4.1	13.1
	Diapason	4.1	11.6
	Oboe	3.5	10.8
	Piano (Trial 1)	5.5	
	Experiment (Aver.)	4.1	15.1

not carried as far as was originally intended. The results from three of the observers are given in Table XIV. The average error per note differs but little from the average error for piano notes. The flute stop has a slightly larger amount of error than that found in the first trial on the piano, whereas the other stops have for the most part less error than that of



trial 1 on the piano and slightly more error than the average error of the practice series on the piano. All the observers when tested on the organ had the advantage of three months' practice in judging notes on the piano. Since the range of notes is smaller for the organ stops than for the piano, the observers had less chance for making gross errors with organ notes than with piano notes. This factor must be considered in comparing their results. The average time per notes for the four organ stops does not differ to any considerable degree from the average time needed to judge piano notes.

e. *Violin errors and correct notes* Co's progress in the violin practice series is similar to that of the piano series,—an irregular decrease in the size of the average error per trial together with an irregular increase in the number of correct notes. C and G were tested a single time in identifying the violin notes one year after the close of their long practice series on the piano. It is significant that but one octave error occurs in any of the results from these tests on the violin. Co often mentioned that the notes on the *A* and the *D* strings

TABLE XV

Records of Errors and Correct Notes of Three Observers in Judging 33 Violin Tones (*g*<sup>#</sup>—*d*<sup>1</sup>).

Observers	Co.												G	C
Trials	1	2	3	4	5	6	7	8	9	10	11		1	1
Size of Error Measured in Semitones	103	79	69	61	36	67	36	37	59	21	31		8	80
Number of Correct Notes	3	2	9	4	15	9	14	10	14	20	16		22	1

seemed alike, but never that octave differences could not be distinguished. Co's interval errors vary from overestimations of a semitone to a minor sixth, and from underestimations of a semitone to a minor seventh. C's interval errors extend to overestimation of a major third, plus one octave error, and to underestimations of a major sixth. G's interval errors are two semitone overestimations and four semitone and one whole



tone underestimation.<sup>8</sup> In these experiments the accuracy of the pitch of the notes depends on the experimenter's keenness of hearing and technique with the instrument. Slight inaccuracies necessarily occur, which make these results in a measure less reliable than those from the study of the piano keyboard.

f. *Kymograph records* Two kymograph records measuring the breathing curve, movements of the larynx, movements of the tongue and changes in volume in the arm were taken during the tests for judgments of violin notes for C and G (see Figures 11, 12, pages 66, 67).<sup>9</sup> During the early part of the experiment C's records show little movement of the larynx or the tongue, and fairly regular although shallow breathing. The breath was usually suspended on hearing the sound until the judgment was made, as is characteristic during tense attention. Later on in the test more movement of the tongue and the larynx is found and the breathing curve becomes increasingly irregular. Absence of any gross movement, except at times during the initial adjusting period characterizes the interval for judgment.

G's curves show more general movement than those of C. The irregular crests of the erratic breathing curve with its periods of held breath indicates also very strained attention (due in part to the observer's interest in the production of the tones as well as in her own task). The curve for movements of the larynx shows usually a gradual depression during the presentation of the note, indicating a lessening of pressure. The curve for tongue movement shows a fairly uniformly slight rise during the judging period, the result of pressure. On the whole there is comparative quiet during the interval between presentation and response in contrast to the continual slight movements recorded at other times. Kinæsthesia, was, therefore, a negligible factor in this experiment in identifying notes. The average time for judgment was 12 seconds for C and 6 seconds for G, which corresponds closely with the average judging time for piano notes.

---

<sup>8</sup>G has always noticed the sympathetic vibration of the open string when that note is played on the adjacent string in one of the higher positions. This would invariably give a clue to the names of certain notes. The other observers apparently made no use of this factor in their identifications.

<sup>9</sup>The plethysmographic records were discarded because they did not give sufficiently clear readings.

g. *Results from the experiment with chords* (See Table XVI) The practice series of observers Bd and G for judging the fundamental tone of the twelve tonic chords based on each note within the middle octave show an initial relatively low ability and fairly regular improvement. The most frequent errors for Bd are in order: semitone, whole tone, minor third, major third, fifth, fourth and augmented fourth; for G:

TABLE XVI

Results for Three Observers in judging the Twelve Tonic Chords played Twice at each Trial on the Piano within the Range  $c^1$ - $c^2$ .

	Bd				G					Hr
	Trials				Trials					Trial
	1	2	3	4	1	2	3	4	5	1
Total Correct	6	7	8	8	9	7	10	14	15	22
Size of Error in Semitones	81	26	28	28	21	20	21	15	11	2

semitone, whole tone, minor third and fourth. H's errors consisted in two semitone errors. Only Bd responded at any time with names of other notes contained in the chord, and only in trial 1 do responses of the fifth occur. Here as in the tests with single notes the errors on the whole were less frequent the larger the interval. Both Hr and G did relatively poorer in this test than in the tests for judgments of single piano tones in the same tonal region. Only two semitone errors out of 84 responses occurred in Hr's judgments in the middle octave and G's errors in this region, out of 276 responses, seldom exceeded a semitone. This is contrary to the views expressed by von Kries, Abraham, Boggs and Köhler. Stumpf reports a similar case.<sup>10</sup> All of the observers in this test had had sufficient musical education to make chord analysis comparatively simple. Some new factor of fusion was entering in to complicate this situation rendering the new tonal blends baffling. That these difficulties could be met by continued practice is evidenced by the improvement in the successive trials.

h. *Results from the experiments with tuning forks* A certain amount of ability in identifying the ten increment

<sup>10</sup>Stumpf, op. cit., vol. 2, pp. 553-554.

and standard (435 v.d.) tuning forks was shown by G (Table XVII). The results for the higher pitched forks are well beyond the limits of pure chance. When the increment from fork to fork is very small the judgments seem little better than guesses. The number of times each fork was responded shows a preference for naming the 3 v. d. and 5 v. d. increment forks. This may be due to the fact that the standard fork (435 v. d.) was always thought of as  $a^1$ , which has a somewhat shifting value,  $a^1$  at concert pitch being 440 v. d. and 435 v. d. for ordinary purposes. Hence the 3 v. d. and 5 v. d. increment forks might on this account be confused at times with the standard fork.

i. *Tonoscope records* The ability to reproduce a tone depends to as large a degree on one's ability to sing the note that one is imagining as on one's accurate formation of an image. False attempts are confusing and interfere with the voluntary production of adequate images. The nine observers who took part in singing the notes  $c^1$  and  $a^1$  into the tonoscope were more successful as a group with  $c^1$  than with  $a^1$ , both in regard to the average vibration rates and the average deviations (Table XVIII). The three violinists F, B, and G, and Co who had completed the practice series for the violin, have the best results of the group for  $a^1$ . It is natural that  $a^1$ , the note to which violins are tuned, should persevere with some accuracy in their minds. The large amount of underestimation in singing  $a^1$  is undoubtedly due to the increased muscular strain in producing higher notes. The largest amount of error for  $c^1$  for all but one observer, A, does not exceed a whole tone; for  $a^1$ , with the same exception, is nearly a major third. Observer A had the poorest record of any member of Group 1 who took part in identifying piano notes. The relatively small errors of the other observers do not offer much ground for comparison with their results in the tests in judging piano notes. The variations in standard pitch of the different pianos make it difficult to arbitrarily select 435 v. d. or 440 v. d. as the basis for grading. Generally speaking in this test the best observers have also been those who excelled in the tone identifications, and the poorest observers those who were relatively less successful in judging piano notes.



TABLE XVII

Showing G's Results in judging Eleven different Tuning Forks Presented in random Order, One a Day until Each had been identified Six Times.

	Standard Tuning Fork		Increment Forks									
	435 v.d.	0.5 v.d.	1 v.d.	2 v.d.	3 v.d.	5 v.d.	8 v.d.	12 v.d.	17 v.d.	23 v.d.	30 v.d.	
Average from six	2.6 v.d.	1.8 v.d.	1.8 v.d.	3.2 v.d.	4.6 v.d.	1.9 v.d.	5.7 v.d.	10.3 v.d.	14.3 v.d.	19.8 v.d.	24.3 v.d.	
Judgments	1.8 v.d.	1.6 v.d.	1.5 v.d.	1.2 v.d.	2.9 v.d.	1.8 v.d.	3.7 v.d.	5.0 v.d.	6.7 v.d.	6.0 v.d.	3.8 v.d.	
Aver. Dev.	2.6 v.d.	1.6 v.d.	1.5 v.d.	1.5 v.d.	3.4 v.d.	2.0 v.d.	3.7 v.d.	5.0 v.d.	6.7 v.d.	6.8 v.d.	5.7 v.d.	
Number of Correct	1	1	1	1	0	1	2	1	1	1	2	
Responses												
Number of Times Each Fork was named	9	4	2	4	11	11	4	7	4	6	4	



TABLE XVIII.

Showing the Vibration Rates averaged from Three Records taken at different Times on the Tonoscope of Tones which the Observers, unaided by any Artificial Devices, believed to be  $c^1$  and  $a^1$ .

Observers	$c^1 = 261$ v.d.			$a^1 = 435$ v.d.		
	Averages	Deviation from 261 v.d.	A.D.	Averages	Deviation from 435 v.d.	A.D.
C	263 v.d.	+2 v.d.	7.0 v.d.	358 v.d.	-77 v.d.	2.0 v.d.
F	263	+2	7.8	418	-17	4.5
G	267	+6	4.3	450	+15	26.0
E	252	-9	1.3	510	+75	5.7
B	285	+24	2.3	408	+27	12.7
D	237	-24	5.0	356	-79	13.3
Co	265	+4	39.7	421	-14	4.7
W	285	+24	25.7	392	-43	17.7
A	404	+143	9.7	579	+144	43.5

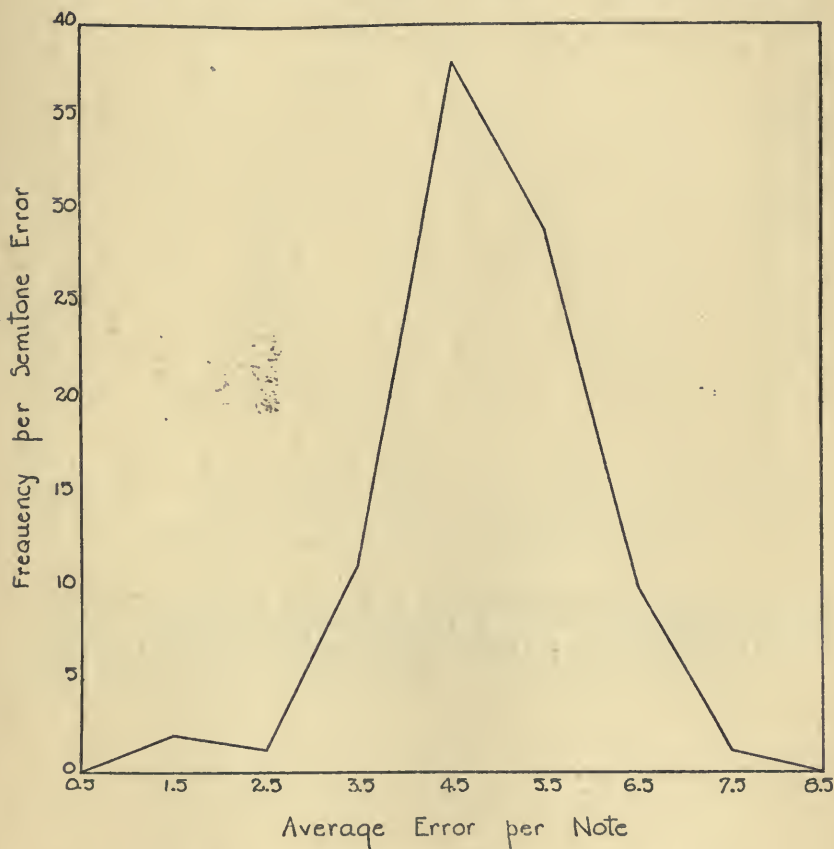


Figure 1.—Frequency-Distribution of Error per Note measured in Semitones and averaged from Total Number of Trials from Groups 1, 2 and 3. (Table I)

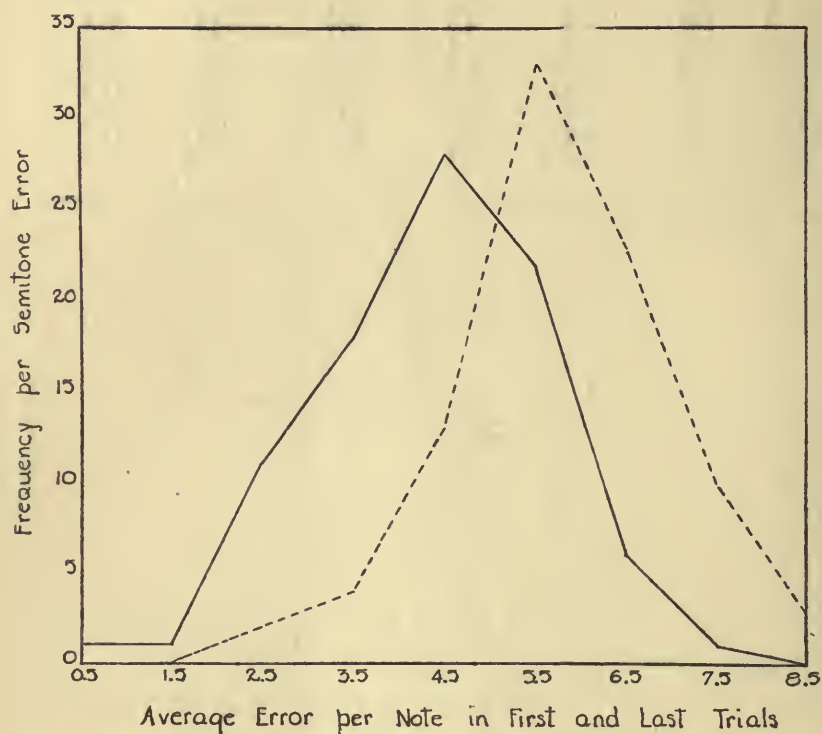


Figure 2.—Frequency-Distribution of Error per Note measured in Semitones from First and Last Trials of the Practice Series. First Trial in Broken Line. Last Trial in Continuous Line. (Tables II and III)

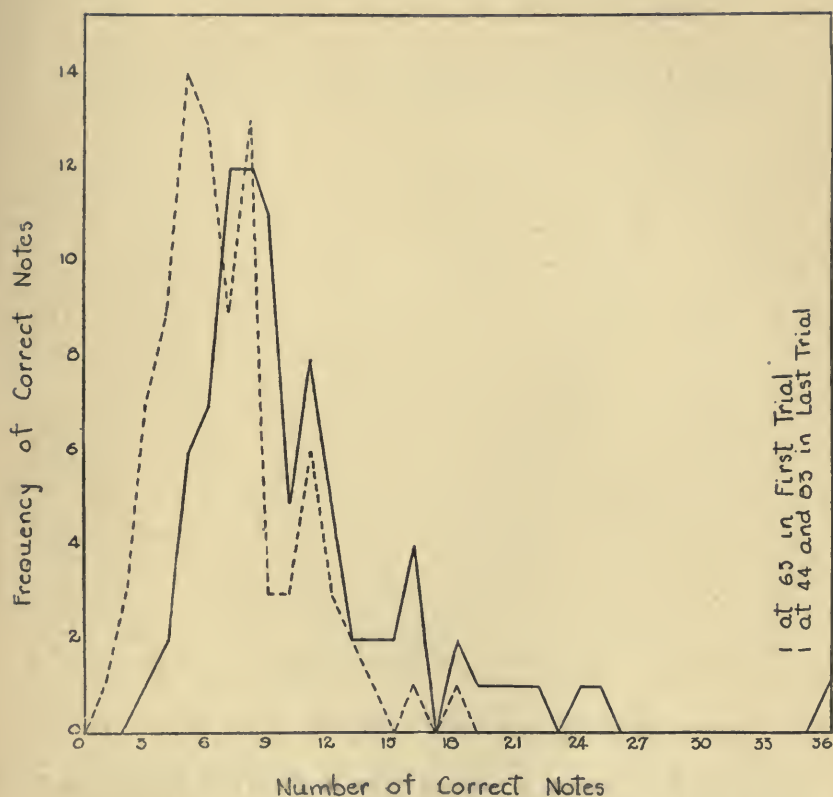


Figure 3.—Frequency-Distribution of Correct Notes averaged from Total Number of Trials from Groups 2 (Continuous Line), 1, added to Group 2 (Dotted Line), and 3, added to Groups 1 and 2 (Broken Line).

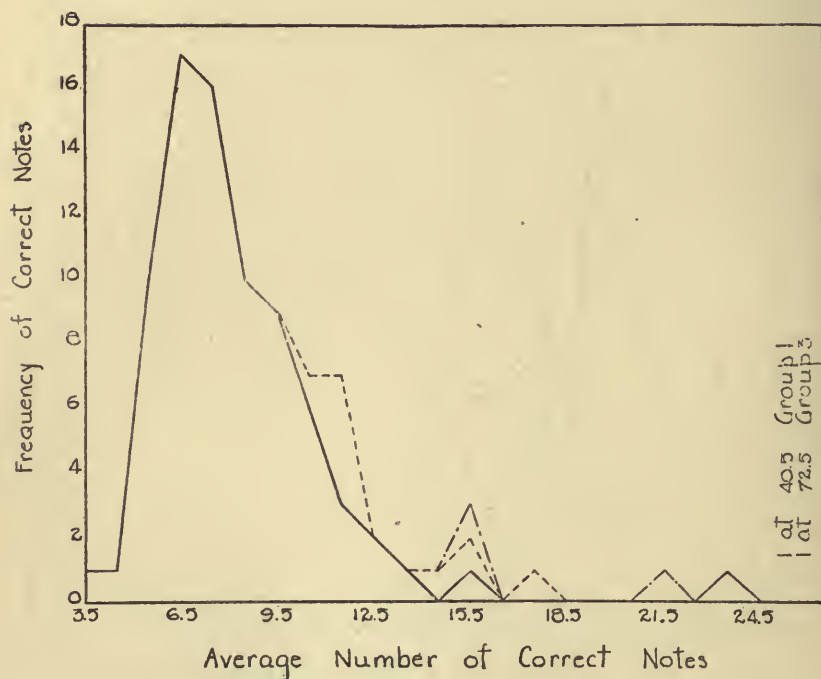


Figure 4.—Distribution of Correct Notes in First and Last Trials of the Practice Series. First Trial in Broken Line, Last Trial in Continuous Line. (Tables V and VI)



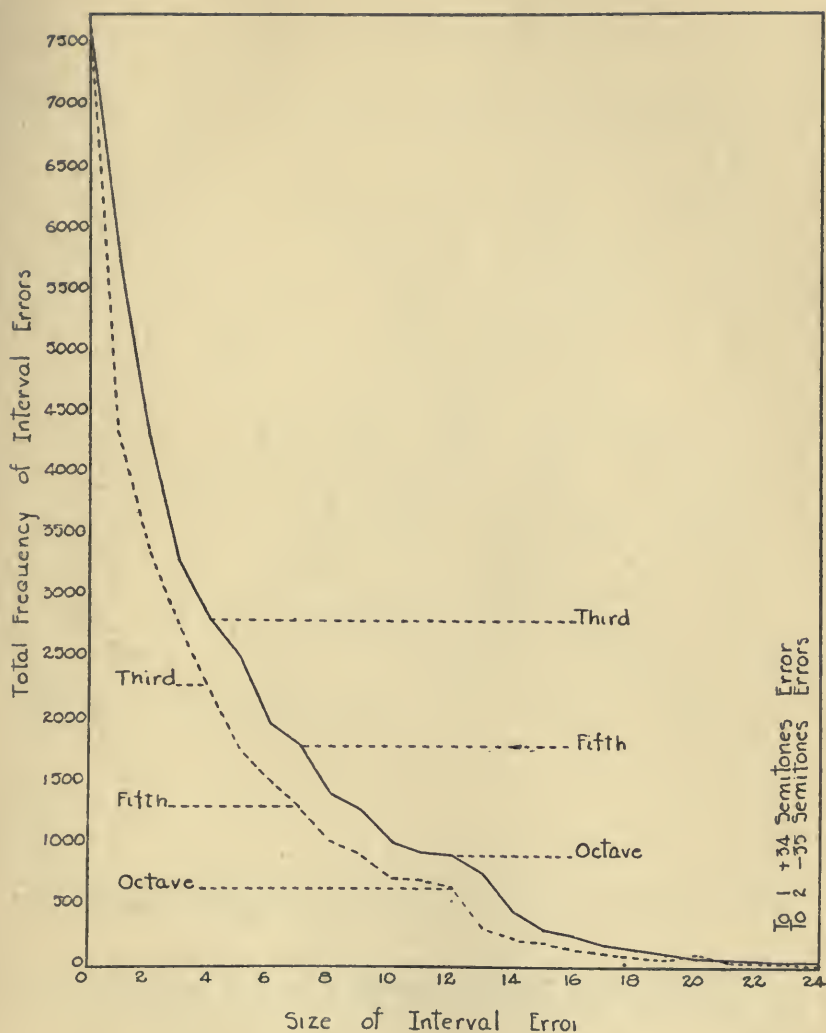


Figure 5.—Total Frequency of Interval Errors obtained from all Records. Broken Line represents Underestimation of Interval, Continuous Line Overestimation of Interval. (Table VII)

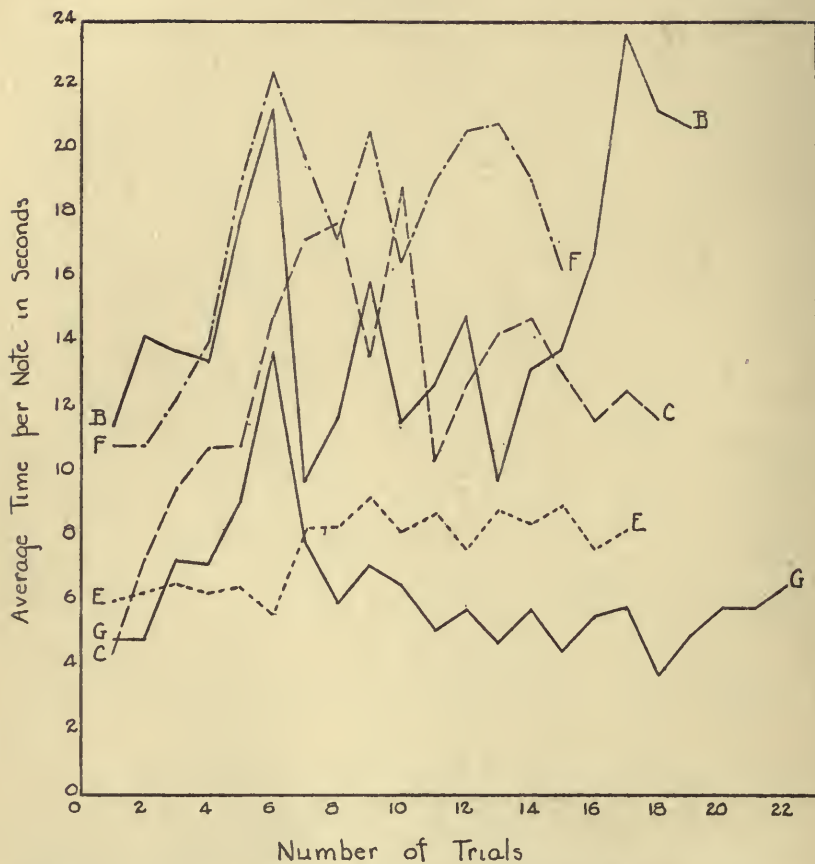


Figure 6.—Average Time per Note in each Trial of the Practice Series for 5 Observers. (Table VIII)

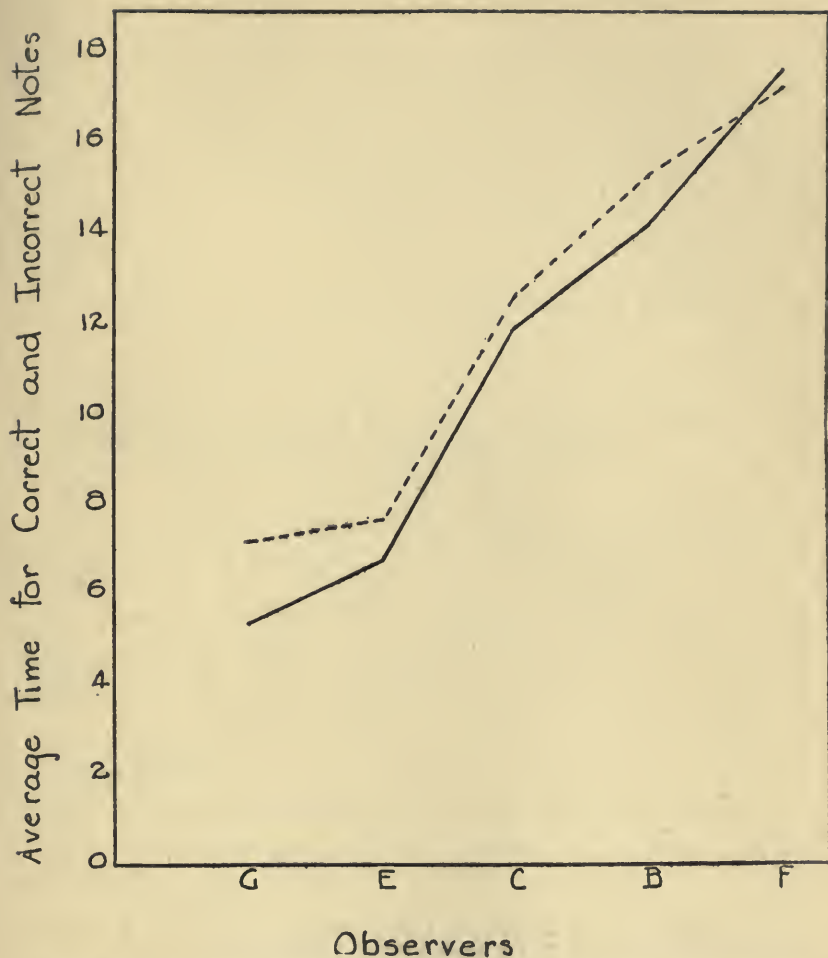


Figure 7.—Average Time for Correct and Incorrect Notes in Seconds. Continuous Line represents Time for Correct Notes, Broken Line Time for Incorrect Notes. 5 observers. (Table IX)

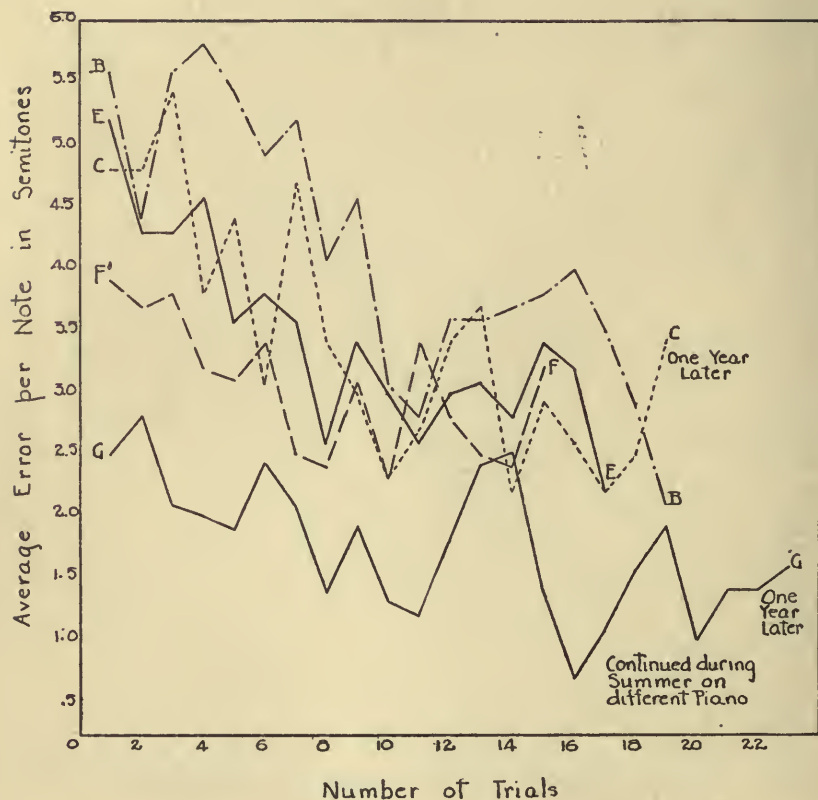


Figure 8.—Practice Curves for Observers B, C, E, F and G, based on Average Error per Note for each Trial. (Table X)



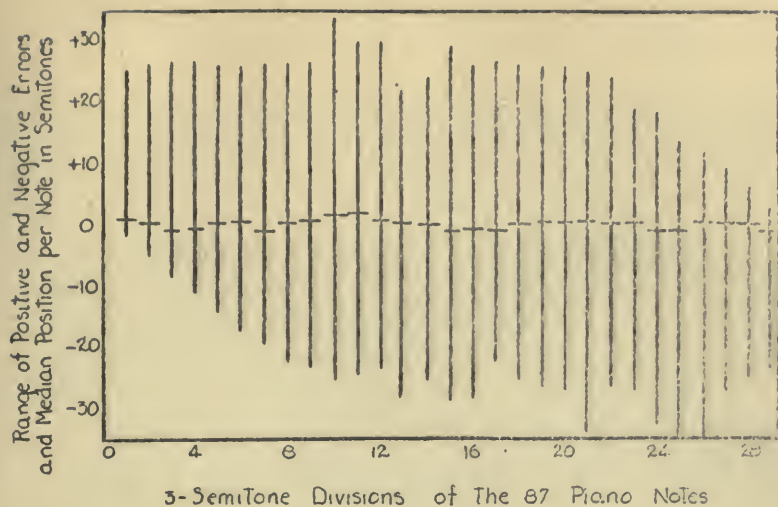


Figure 9.—Range of Positive and Negative Errors and Median Position for each 3-Semitone Division of the 87 Piano Notes.

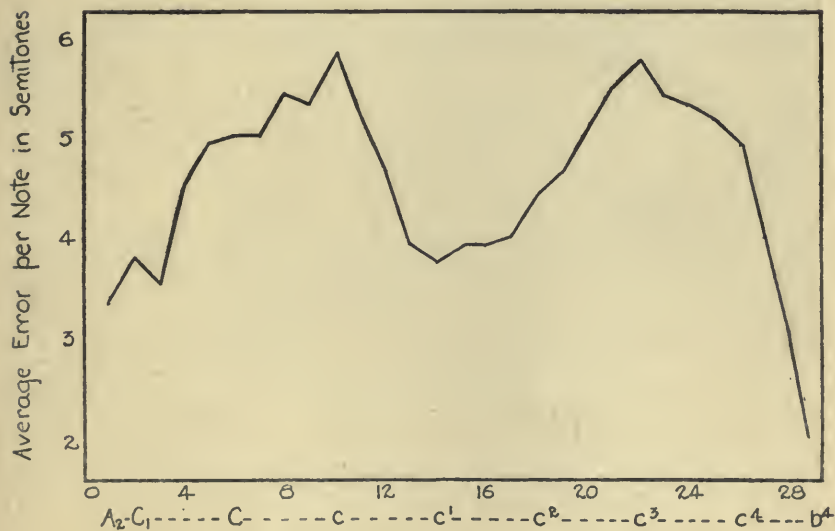


Figure 10.—Average Error per Note per 3-Semitone Division based on 725 Trials from Groups 1, 2 and 3. (Table XIII)

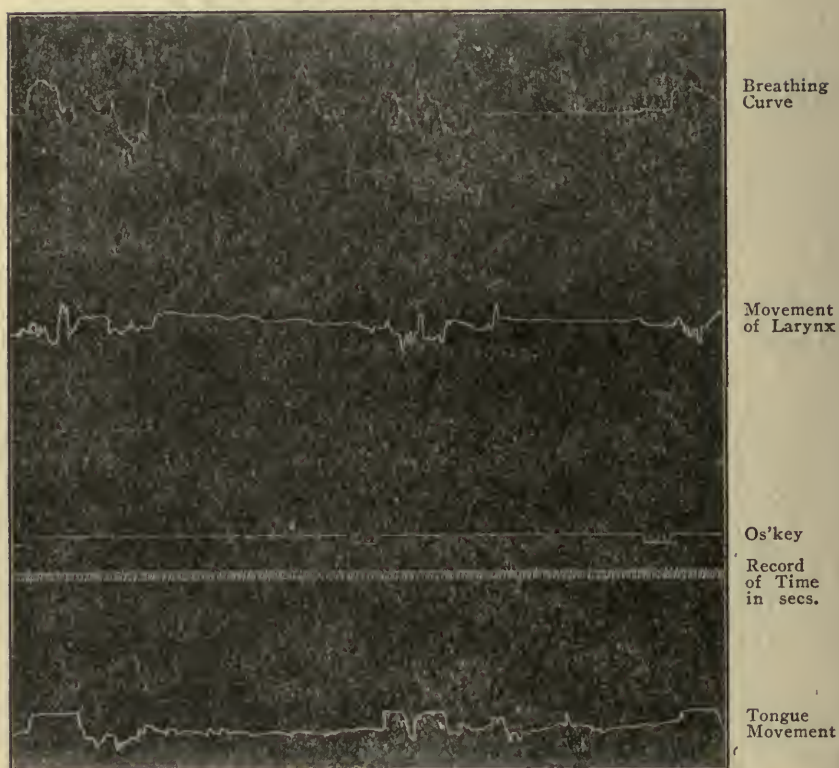


Figure 11.—Kymograph Records of Movement during Period of Judging Violin Notes by Observer G.

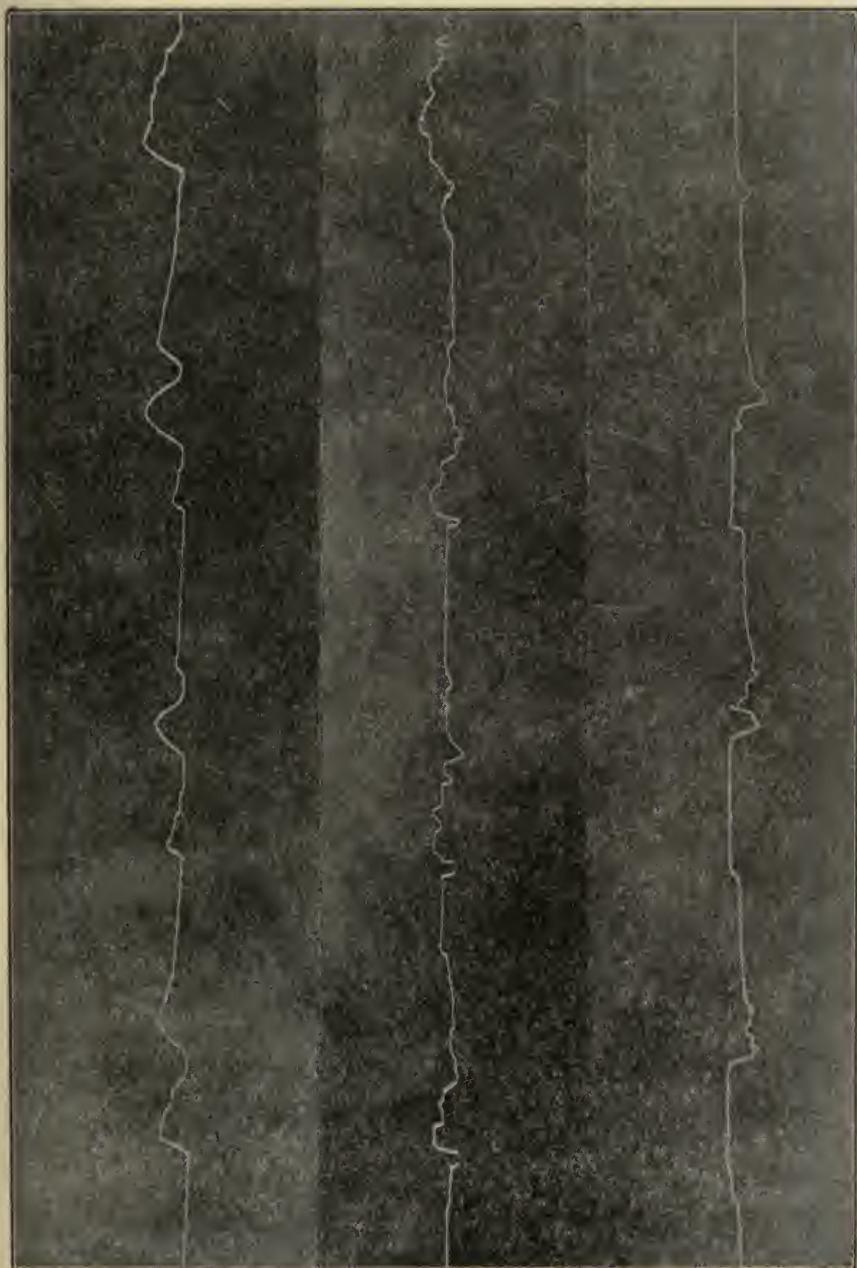
Breathing  
CurveMovement  
of LarynxTongue  
Movement

Figure 12.—Kymograph Record of Movement during Period of Judging Violin Notes by Observer C.

## V.

### Analysis of Introspective Reports

The responses to the questionnaire, which are given in full on pages 30-33, bring out many striking features. First, the observers in Groups 1 and 2 have had distinct musical advantages. Only four observers had no practical music training of any kind. A summary of the frequencies of individual ratings for musical ability on a basis of 0 to 10, from 74 observers is as follows:

Ratings	0	1	2	3	4	5	6	7	8	9	10
Frequencies	7	2	4	6	12	14	6	14	7	2	0

The frequencies of rating for musical appreciation are as follows:

Ratings	0	1	2	3	4	5	6	7	8	9	10
Frequencies	1	0	1	1	3	13	9	15	19	9	3

The median frequency for musical ability falls on grade 5, and for musical appreciation on grade 7. The frequencies for individual estimates of good, fair and poor in musical memory are as follows:

Memory	Good	Fair	Poor
Frequencies	22	38	14

The tests for imagery yield the following results:

Most Vivid Type of Imagery	Visual	Auditory	Motor	Mixed
Frequencies	48	10	8	8

Half of the observers who possess a predominance of auditory imagery rank very near the top in ability to identify notes, the others are scattered throughout the middle ranks. None with good auditory imagery are found near the poor extreme in this ability. No similar correspondence is found for the other types.

The literature on tonal identification assumes that people with marked ability in this respect always recognize first the note-name, as *c* or *d*, then after momentary reflection give it a place in an octave, as *c*<sup>3</sup> or *d*<sup>4</sup>. The questionnaire reveals



the fact that fourteen observers were accustomed to identifying the note directly, not for example as some *d* or other, but as *d*<sup>1</sup>. The observers who made this a single process include some of the most able in identification and others less able. Seven of the 79 reporting recognize the names of the note first and later its octave. Three of these seven were notably good in their judgments, the other four only fair. Forty-six observers named the octave first and then gave the note. Twelve occasionally named the note first, occasionally the octave first.

In reporting direct or indirect judgments twenty-six usually gave the immediate or direct judgment, thirty-four usually delayed judgment for consideration, ten at some times gave direct judgments at other times indirect judgments.

Thirty-eight preferred strongly to judge the note from its initial sound, immediately after it was played (which necessitated many short repetitions of the note), twenty-five preferred to have it sound a few moments before trying to make a judgment, and two were impartial. Those who judged from the initial sound had on the whole the best results.

In response to the question regarding the setting which each note was given, forty-five gave the note a spatial position on either an imagined piano keyboard or on the cardboard chart placed before them. Seven thought of it as belonging to a particular scale, nine as belonging to a chord, one as the fifth of a chord, others as the fundamental of a tonic chord, and others as the tonic or the dominant seventh of a chord followed by its proper resolution, ten combined a number of the settings already mentioned, and five claimed to have no definite namable setting for the note. Several referred the note frequently to imagined violin notes, and one had an imaginative ladder scheme which represented the ascending and the descending scales. Those who gave the notes definite settings in chords had on the whole the best objective results.

The aids to recognition which are enumerated in the reports are, (1) kinaesthetic sensations (vocal motor and other forms), (2) knowledge of the vocal range, (3) auditory memory of the individual notes assigned, (4) auditory images of the tonic chord or other chord combinations with harmonic resolutions, (5) auditory images of notes a third or a fifth or an octave above the note presented, (6) auditory memory of the violin *a*<sup>1</sup> and other notes, and the piano *c*<sup>1</sup>, and in gener-

al the *c*'s, *b*'s, and *f*'s on the piano, (7) recollections of certain pieces, (8) recognition of certain tonal characteristics, and (9) tonal volume. The hindrances to recognition enumerated in the reports are, (1) a shifting standard of some pitch or pitches in memory, (2) distracting noises (including distant music, occasional faint humming, not permitted in the group, the sounds of lawn mowers and bursts of song from nearby birds, which could not at all times be eliminated), (3) wavering of the attention through fatigue or distraction, (4) the unfamiliar tone-quality of the piano, and (5) the inability to produce voluntarily auditory images.

The majority of the observers regarded the first part of the sitting as the easiest because then they felt fresh and could control their attention better, and the last part of the sitting the hardest, because as the experiment continued their imaginations flagged increasingly through fatigue and distraction. A few observers felt that the task grew easier during the course of the sitting, as they overcame the first strangeness of the situation and adapted themselves to a kind of temporary auditory set which facilitated judgment.

Twenty-seven observers report that they found the middle range of the piano the most difficult to judge and the extreme octaves the easiest. Twenty-five observers found the middle range the easiest and the low notes hardest, and eight observers thought the reverse to be the case. (See Table XII, page 48). In explanation many said that the middle octaves were more familiar to them on account of the frequency with which they had heard them, and on account of their ability to form auditory kinaesthetic images of notes in this range. On the other hand, many said this range was the hardest for them to judge because the character of all the sounds was very much the same. The extremes were claimed to be difficult because of the lack of definite pitch there. Others noticed distinguishable characteristics at the extremes not found in the middle of the keyboard. The fact that within limits underestimations could not be made at the lower end of the keyboard or overestimations at the upper end eliminated a source of error found in the middle range.

Forty-five observers expressed themselves as having enjoyed the experiment throughout its course. They were interested in watching their progress and anticipating results.

Fifteen found it very difficult and tedious. Eleven claimed to have maintained a neutral attitude.

A survey of the introspections recorded by the observers at each trial throws light on some of the problems involved in tonal recognition and possible methods of meeting them.<sup>1</sup> It is striking how few allusions are made directly to the pitch of the notes. Several observers write: "Tried to get the general range and tonal qualities of the different piano notes. The quality gets deeper and more resonant as the notes go down. Each octave has a different quality of its own." "After practice noticed that the different notes seem to have different volumes." "The volume of the note helps to place it in the right octave." "The last octave has a certain metallic ring." "Prefer low notes because they jar less than high ones and can be attended to better." "The difference between octaves is one of intensity." (This observer claims to be tone-deaf.) "Tone color helped." "The higher notes do not last (resound) as long as the lower ones. Many notes almost pained the ear." "Listened for the vibrating quality which differentiates the high notes from the lower. High notes are thin, the low ones deep and full." "Determined approximate position by volume."

"Nearly always judge the notes in relation to whole keyboard rather than to a particular note. Have a feeling of spacing. There is a certain point at which the note must lie." "A sense of position seems to be the only guide. I hear the sound and then by reasoning place it where it seems to belong on the keyboard. Octaves help little. I have simply a notion of the total position."

A few of the difficulties encountered were: "Octaves at the center of the piano are harder to distinguish than those at the ends." "Seem to be a great many notes in the *small* octave." "After listening to a number of notes, it sounds as if the same notes were repeated many times, especially those near the middle of the piano." "Too many notes between  $c$  and  $c^2$ . Do not know what to do with them all. Forget

---

<sup>1</sup>No attempt to direct the course of the introspections by suggestion was made at any time. As was stated, these reports were read each week and criticized in regard to clearness of expression and completeness, in order to stimulate the observers to attentive analysis. All the observers had used Titchener's *A Beginner's Psychology* and were familiar with his terminology in the chapter on Sensation.



that the octave  $c^2$  to  $c^3$  exists." "Practice seems to make many notes sound alike."

For the most part practice tended toward impressing octave differences on the observers. "Practice makes it easier to decide on the octave than before, but it is just as difficult to find the right note in the octave." (This observer had the poorest record of all.)

Various methods were employed to overcome these difficulties. "Tried to recall  $c^1$  and judge from that." "At first tried to refer all notes to  $c^1$ ." "Judge from memory of the notes in the middle octave. The lower and higher notes seem to resolve into one of the notes between  $c^1$  and  $c^2$ , not when first sounded, but after the second or third striking. Judge better by listening for overtones. Keep in memory the tone quality of  $c$ ,  $e$  and  $g$ ." "If I thought the note sounded like  $f^\sharp$ , I imagined  $g$  as following it and estimated the probability."

"The first, third and fifth notes in the octave have a more finished tone than the others, which helps in locating the note." "Some notes have a familiar sound, but I cannot always tell whether it is a  $c$  or an  $f$ ." "The notes begin to have a definite letter attached to them when they are struck. I do not have to hunt for a place on the chart as was necessary at first. Something familiar in the quality of  $c$ ." "This time I did not have to think so long what note it might be. A name just seemed to come into my mind."

The members of Group 1 had more opportunity to attempt and develop new methods in their longer series of practice. Quotations from their weekly introspective reports show their progress.

Observer Co, practice series for the violin.

Trial 1. "No particular method. All guesses, I think."

Trial 2. "Still guesses. Tried to use relative pitch without much success."

Trial 3. "The  $c$ 's are getting easier to identify, also the  $f$ 's which have a 'squeezed in' compressed quality as a rule."

Trial 7. "Several times I found myself having a feeling of familiarity when a certain note was struck. Usually I thought it was  $c$ -or  $a$ , or  $g$  or  $b$ , and then had to decide, sometimes by humming, which of the two it was. With the other notes, I identified the string on which they were being played and then the position of the note."

Trial 9. "Although I succeeded in confusing the  $d$  and  $a$



strings to some extent, the notes on these two seemed to possess something of a 'quale'."

Trial 10. "Usual method now is either (1) immediate or almost immediate recognition, after which I usually try to verify; or (2) visual image of string and kinaesthetic imagery with little auditory; or (3) absolute blank, followed by an attempt to follow out procedure (2)."

Trial 11. "It is seldom a note is absolutely certain in my mind upon hearing it. Usually I have to go through a series of checks after I have decided what it is. In general these checks are kinaesthesia (vocal motor sensations), related auditory images and a few visual images in connection with the auditory."

Same observer Co, practice series on the piano, following the series on the violin.

Trial 1. "All the octaves save the middle one and a few notes on either side were guesses. With the notes of the middle octave, I used kinaesthetic sensations and actual humming to identify them. The other notes I tried to relate to the corresponding notes in the middle octave."

Trial 2. "Related the notes to those on the violin. *G* seems to me to be the lowest note which has real 'roundness'."

Trial 4. "Tried to do away with the humming in my judgments this time. Took my snap judgment and then tried to find out the peculiar 'quality' of that particular note, and fitted them together. If they did not fit, I had to discard my snap judgment and try the humming method or else let it go as a mere guess."

Trial 6. "The notes seemed to have 'qualities' for me for the first time. It seemed as though I could have finally distinguished many of them if I had taken time to associate that 'quality' with the same note in the middle octave."

Observer B, practice series on the piano.

Trial 1. "Tried at first to relate the notes to a remembered  $c^1$ . Found I could not remember  $c^1$ , when another note was played. Noticed I could remember the violin  $a^1$ , so related notes to it. Related the extreme octaves to the middle octave."

Trial 2. "When a note was struck I hummed it in the middle octave calling it *do*, and hummed the *me, sol, do* above."

Trial 13. "The following factors help me in naming a note, (1) humming the note and its diatonic scale. If the

note is not within my range, I give it a position in the scale I can hum, and later place it in its right octave. (2) I seem to tell whether the note is a sharp or a natural by the way it is struck. (3) The highest and lowest notes I tell by the amount of tone."

Observer E, practice series on the piano.

Trial 1. "Felt absolutely helpless when I first heard the notes. Could not tell one octave from the other. I had an idea of  $c^1$ , but could not tell when it was heard."

Trial 2. "Felt a little surer of myself. During the practice periods of this week I spent most of the time on the middle, the second and the highest octaves, trying to detect slight differences between notes. Certain notes like  $e$  and  $f$  seem to have the same quality, which make them difficult to distinguish. I gave up the idea of singing the notes and attempted to tell directly by their sound. This seems to be more satisfactory."

Trial 4. "Notice that the thrice-accented octave has a marked difference in quality from the others. It is more rasping."

Trial 10. "The placing of the octave does not bother me a great deal. I felt fairly sure of the octaves except the two lowest."

Trial 11. "The middle octaves I sing to be sure of the notes. Others I judge entirely by their pitch. By having the note repeated if I am not sure at first, I can get an immediate judgment or nearly that."

Trial 12. "Feel that I am improving. The quality of the note rather than the pitch locates it for me in the scale."

Observer F, practice series for the piano.

Trial 5. "When I started this trial I had an auditory image of  $a^1$  firmly in mind, as a result of previous practice. The first few notes were correctly named through relative pitch. Then a low note threw me off. A great deal of the time I first listened carefully to the sound of the note and tried to think what note it was without using indirect means, and then compared the result with a muscular sensation in the throat. Some times no comparison was necessary, for I was certain what note it was without any comparison. Whenever I came across one of these notes of which I felt certain, which were  $a$ 's,  $d$ 's,  $e$ 's, and  $c$ 's, I used relative pitch with the

succeeding notes until some wide interval skip between notes put me off."

Trial 6. "Depended mostly on humming the note in the singing range. Sometimes just listened to the notes. Got more right by the latter method."

Trial 10. "The notes now seem to have different qualities,—some are open, some vibrate more than others and some are dull. They also seem to have more place on the scale.

Trial 14. "It is very difficult for me to keep my attention fixed. My ears feel strained listening for overtones or purity of tone. I recognize the sound of some of the notes and then forget the name to be associated with it. It seems an interminable task to differentiate between so many notes." (F had twenty-five correct judgments in this trial.)

Characteristics which F associated with some of the notes:

- C real tone not heard.
- G funny overtone wave.
- D weaker after it is held a few seconds.
- E<sub>b</sub> very strong.
- E not as strong as E<sub>b</sub>.
- A buzz or low rattle—hard.
- B<sub>b</sub> thump when struck.
- B funny sound when key is lifted.
- c flats after a second.
- c<sup>#</sup> second overtone rings out.
- d flats when struck—strings not tuned together.
- e<sub>b</sub> deep, hollow sound.
- c<sup>1</sup> flat—no overtones—harsh.
- c<sup>#1</sup> after a few seconds part of tone flats.
- f<sup>#1</sup> rattle when played softly.
- g<sup>#1</sup> thump.
- a<sup>1</sup> open, rings out, not as much as e<sup>2</sup>.
- d<sup>2</sup> rings out but softer.
- d<sup>#2</sup> hollow when key is lifted.
- e<sup>2</sup> rings out more than c<sup>2</sup>.
- g<sup>4</sup> sharp as a knife.
- a<sup>4</sup> very sharp and clear.
- b<sub>b</sub><sup>4</sup> muffled.
- b<sup>4</sup> woody, longer sound than c<sup>5</sup>.
- c<sup>5</sup> woody.



Observer C, practice series for the piano.

Trial 1. "A feeling of amusement at my helplessness in such a situation. However, a strong determination to do as well as possible. When seated before the scale, listening for the note, my eyes were resting upon the middle of the scale. When the note was struck my eyes turned either to the right or left, according to whether the note sounded high or low, and continued to move until they seemed to reach the note heard. There was no conscious analysis of tones. But my idea of the position of the note was so definite in mind that my eyes readily stopped at some fixed place, and the judgments were made with slight hesitation. This made me conclude that discriminations need not be due to conscious comparison, but that a method of training to develop immediate judgment would be adequate."

Trial 2. "During period of practice I sounded the various C's on the piano in order to get a clear idea of the characteristics of the different octaves. I noticed when the low notes began to have distinct overtones and where the high notes disappeared into their more prominent overtones. I tried to differentiate the timbre of the lowest two bass octaves and also that of the highest two octaves. At the end of the practice I felt confident that at least I could place a note in its proper octave.

"This test was made after my partner had been tested. She used a method of conscious comparison, humming to get the interval more clear in mind. Her results were on the whole quite accurate. I realized that that method was probably more reliable than mine, so I undertook my own trial with some hesitation. At first I tried to use her method, but realized that I had not had sufficient training for it. I then resorted to the method of immediate judgment which had previously appealed to me. I found I was reduced largely to guessing and that my practice to try and differentiate the octaves had amounted to little. The lowest and highest notes were not as distinct from each other as I had anticipated. Feeling of discouragement."

Trial 3. "During the practice period I endeavored to learn to differentiate the octaves, as before. I tried to analyze the quality of each, so that I could at least place the note in the proper octave. More attempt to get the note placed in the right octave, than to get the correct name of the note. The



name of the note usually came readily to mind without conscious comparison, but the determination of its octave was still perplexing. Feeling of helplessness usually with very little confidence, especially in regard to the high and low notes."

Trial 4. "During the practice period I was anxious to be able to name the note correctly as well as place it in the octave. I played over the middle octave many times trying to fix it in mind, in pitch and relative variations in pitch. When I found I could hum it readily and accurately, I hit at random notes on piano to see if I could name them correctly. Result—small percentage of right answers. Used method of conscious comparison. I tried to place the note heard in the scale of  $c^1$ - $c^2$  and then place it in the proper octave. My reaction took longer but I had more of a feeling of confidence. There was difficulty in being able to keep pitch of  $c^1$  very clearly in mind. I did not attempt to locate sharps or flats to any extent. Not entirely satisfied with this method."

Trial 5. "During the practice period I tried to fix separate intervals clearly in my mind. From the sound of middle  $c$  I would try to work out the pitch of notes at various intervals from this. I sang the notes, trying to train my ear to become more accurate. After I had secured the note I played it on the piano, in order to see how correct my judgment had been. There was great inaccuracy at first but I found improvement before the practice was completed. The pitch was determined by singing *do mi sol do* from middle  $c$  to try and locate the given tone. I still felt very hesitant about my ability and I could not see in what way my practice had helped me. During this sitting it occurred to me that I could use my highest and lowest singing notes for standards which would be less likely to vary than my memory of middle  $c$ . I determined to attempt that at my next practice period."

Trial 6. "During the practice period I played the various high and low notes in order to try and hum their octave in middle range. I found that I frequently judged their pitch a fifth or a third off so I continued my practice until I could readily distinguish the octave of the note struck. Then I tried to discover my high and low thresholds for humming tones so that I could use them for bases for comparison. I found I could get the pitch of the given tone more clear in mind than at any previous sitting. But it was still hard work to name

the note. I sang down the scale as far as I could and tried to learn from the last note what scale I had been singing. I felt still very uncertain about my accuracy. In the case of middle *c* I judged it correctly more from the feeling of familiarity accompanying it than from its relation to my lower threshold."

Trial 7. "During the practice period I found that my humming ability at the highest limit varied more than at the lowest limit, so I decided to adopt only *c*, my lowest limit, as my standard. I hummed it over and over, endeavoring to fix in mind the auditory and kinaesthetic sensations. Then I studied out how I could infer the pitch of any note on the piano with reference to this standard. When the note was heard I tried to hum it, or its octave equivalent within my range, then, by means of *do mi sol do*, I tried to relate it to *c*. This method seemed dependable, and, although some notes were hard to determine, on the whole I felt quite confident in my judgments."

Trial 8. "During the practice period I decided that *do mi sol do* was not detailed enough, so I attempted to study out the relation of each note on the piano to *c* in terms of half notes, not larger intervals. I decided to sing every note in the scale, upon approaching *c*, not merely *do mi sol do*. Although this method seemed more accurate, it took more time. Moreover, I felt throughout that my judgment of *c* was not entirely fixed. Therefore, more of my judgments might be wrong than before. The method seemed quite unsatisfactory.

Trial 9. "During the practice period I realized I could not make use of the method of singing every note as it took so long. I therefore decided to go back to the method of *do mi sol do*, still attempting to keep the pitch of *c* more surely in mind. Throughout I felt confident in my judgments. But I did feel, too, that the method of humming was not an ideal one, as it involved too much time and had at best a good deal of variability, as shown in my results."

Trial 10. "During the practice period I struck a few notes to see if I could guess their pitch immediately. I found I guessed a few right, the others being very nearly so. So I decided to abandon my former method of determining pitch by humming and to adopt the new way of judging immediately. I tried to fix in my mind the qualities of the various notes at the top of the piano, analyzing them as carefully as

possible. Afterwards I attempted to do the same with the lower notes but I did not have time to differentiate them as well. It was as interesting as a game. I judged the pitch of each note as accurately and quickly as possible, running my eyes up or down on the chart until they seemed to reach the right spot. On the whole, I felt quite confident, especially with the higher notes. When I was not certain of the octave, I would resort to my old method of humming."

Trial 11. "During the practice period, as before, I tried to differentiate the notes at both extremes of the piano. I also tried to single out significant points, above or below which the notes seemed to have a certain quality. I practiced especially the lowest notes. I tried to judge the pitch of each note immediately as before. Some I felt confident about, but most of them seemed doubtful, especially those in the middle range. I did at times fall back upon my old method of humming."

Trial 12. "During the practice period I practiced especially the notes of the middle range, trying to get the characteristics of the various octaves. I found significant differences, but they did not limit themselves to the octaves. Some octaves were very much alike throughout while others seemed to have several distinct differences within them. I made my judgments with no great feeling of confidence. But I had become convinced that this was the better method to follow and that in time I might make marked improvement."

Observer G, practice series for the piano.

Trial 1. "Began the test with the idea that I could identify a great many notes by immediate judgment, but was surprised by the unusually penetrating, brilliant sound of the piano notes. This together with the high pitch of the piano confused me thoroughly. I felt some degree of certainty about the octaves, although attention was not always steady on this point which resulted in failure to distinguish between the different octaves. I was more certain about the middle range than other parts of the keyboard. At the extremes, I used chiefly guessing, with no foundation at all for my judgments."

Trial 3. "Attention was very strained, so that genuine fatigue entered in. At these times octave differences are apt to be overlooked. As each note was struck I seemed to go through a definite period of orientation. No vivid imagery



of any sort appeared, but fleeting, hazy, concrete visual and auditory images were present. A slight process of comparison and judging occurred frequently. For instance, a very high note was thin, wiry, penetrating and short. That would be judged as belonging to highest octave, its place in the octave depending mostly on its metallic quality. At the extreme upper and lower ends of the piano pitch entered little into the judgments. Within the other five octaves recognition often came instantly, as distinctly a *b* or an *f* quality, though usually I weighed this judgment a little, considering that this piano was tuned high. If it was difficult to tell at once, I regarded the note as the seventh of a dominant chord and resolved it downward or considered it as the seventh in a scale and thought of the note above, or considered it as the fundamental, third, fourth, or fifth of a tonic chord. If this were too prolonged I forgot the tone, or became quite confused. My memory for the note struck seems exceptionally short. In the once-accented octave I try to attach a certain quality to each note, and believe I am partially succeeding. By transferring this octave's relationships to other octaves, I may be able to extend the range of the ability."

Trial 4. "A heavy cold at this trial made me feel that this experiment were progressing more poorly than others. It seemed as if it were impossible to make use of some standard for judgment to which I was accustomed. I had never consciously used kinaesthetic sensations or images in order to judge by the "feel" or by comparing with my singing range but was led to think by this fact that I actually had used one or the other. However, my results were better than usual, so this may or may not have been a factor. When I tried humming the tone given I was thrown off completely. At no time did it seem to help."

Trial 6. "In practicing I played simple scales, usually the scale of *c*, over and over, noticing the octave variations. Then I tried arpeggios from one end of the keyboard to the other, and simple chord progressions. I spent considerable time practicing the first and last octaves. Their pitches are not very obvious to me as I play them, although each note is sufficiently different from the adjacent ones. At the upper extreme everything from  $b^3$  to  $b^4$  sounds like  $e^4$ .  $c^4$ , even when I am playing it myself, has nothing clearly in common with the lower *c*'s on the piano. I can call up at will good auditory



images of the piano notes from *C* through  $a^2$ . Below that I can imagine only a general, rough sound, and above that only a metallic sound without definite pitch relations. Practice seems to be giving these tones slightly more individual character."

Trial 12. "Like to have the note sounded fairly strongly but not pounded. Frequently ask to have it sounded again softer and louder with a pause in between. The initial sound seems especially important and while the note is still sounding, I ask to have it repeated in order to get that initial sound. Try to free my mind from any bias and 'set' in regard to what note is coming. When the note is heard I place it instantly unless it is very high or very low (when I am usually in doubt and make guesses). Then I like to have the note repeated in order to verify my judgment, and think of it mentally as belonging in a particular chord or scale. On some days it seems hard to judge octaves. The bare, undecorated room in the Music Annex reinforces the overtones to a great extent, and I seem to be judging the width (extensity) of the sound to make sure of its octave. Occasionally I hum the octaves—seldom individual notes. On this account, I like the intensity kept constant and the note struck decisively.

"Have come to know certain peculiarities of our piano;  $c^1$  to me sounds a bit flat;  $a^1$  has an individual quality of tone which is almost never mistaken;  $d$ ,  $b^1$ ,  $c^2$ ,  $e^2$ ,  $c^3$  and  $d^3$  seem to be exactly in tune and are always judged correctly when heard and almost never given when not heard.  $D_1$ ,  $b$ ,  $g^1$ ,  $a^2$  and  $b^2$  I am beginning to recognize as having a certain quality. Other notes are placed as 'not being these,' but as having an approximately definite place. Except at the extremes of the piano, and except when attention is distracted, am certain that the note lies within a very small range of possibilities.  $c^{\sharp 2}$  and  $d^2$  do not seem to be in tune, i.e.,  $d^2$  sounds flat, but I am beginning first to judge  $d^2$  correctly. The highest octave is beginning to have recognizable quality of tone, but it takes much more attention and activity on my part than the lower ones. The notes below  $G_1$  I judge by width of tone and roughness, except, I think,  $D_1$  and  $C_1$ .

"Feel as if attention were a very large factor. It is partly attention to discriminable tone-qualities, not wholly pitch. For instance, when  $g^1$  was played, I knew it to be lower than  $g^{\sharp 1}$ , but was not sure whether it should be called  $g^1$  or  $g^{\sharp}$ . I could

not tell at once whether the difference lay in pitch or in volume." (This confusion occurred on several occasions.)

At the close of the trial which occurred one year after the practice series with piano notes the following reports were made by members of the group.

Observers	Number of Notes Correct at this Trial	Reports
1	11	"Judgments were made quickly, based upon an estimate of the octave in which the note lay. Timbre and length of time notes resounds after it is struck aided in this."
2	44	"Several devices remained from the previous practice. Remembered the type of quality of $c^1$ , $c^2$ and $c^3$ . All seemed very clearly $c$ 's this time. No great assurance throughout. Gave note instant location in some octave. Afterwards tested judgment of note to make sure of it. Remembered octaves well. Verbal imagery (suggesting possible responses to myself) came in. Used the chart very little."
3	5	"Remembered that I was apt to omit entirely the octave $c$ to $c^1$ , and tried to place unfamiliar notes there. Remembered peculiar qualities of certain notes."
4	5	"Used last year's method. Auditory images of the successive octaves above or below the one presented to determine position on the keyboard. Much tension in throat for low notes, less for those in middle range."
5	10	"Remembered the peculiar ring of some notes."
6	6	"Could decide more quickly than last year. Referred all notes except those in extreme octaves to $c^1$ ."
7	10	"Seemed easier and quicker than last year. Remembered the low and high notes better than the middle ones."
8	15	"No definite method of discrimination. Some visual, auditory and motor imagery."
9	8	"Decisions made more quickly."
10	19	"Tried first to recognize the octave, then the note. Some notes sounded familiar, others I had to guess at."
11	9	"Last year I thought of the notes in their relation to $c^1$ . This year, I thought of their relation to each other in the scale."
12	21	"I do not remember anything from last year's practice, but in the past year I have found myself constantly trying to tell in what key a piece is being played, and to pick out certain notes by name."

After the tests in identifying the fundamentals of tonic chords the following introspective reports were given:

Observer Bd, "At first tried to pick out the lowest note and sing it to myself; later just listened for the effect of the whole."

Observer G, "Tried to pick out the individual notes from the chord struck and judge from them. This was very difficult to do for the notes were evenly struck and blended well. In a general way I felt a slight emotional (or kinaesthetic) reaction to the chord as a whole. *c e g c* was instantly recognized. *db* I knew as being not *c* but approximately in that region."

Fundamental of chord	Bd's characterization of chord	G's characterization of chord
<i>c</i>	Blended, hollow	Most complete of all
<i>c#</i> or <i>db</i>	Tiresome	Unsatisfied
<i>d</i>	Not clean cut	Wistful, not at all self-sufficient
<i>d#</i> or <i>eb</i>	Hard, ringing	Some independence
<i>e</i>	More clear cut	Cold, impersonal
<i>f</i>	Hollow	Finished effect
<i>f#</i> or <i>gb</i>	Not clear cut	Cold, powerful
<i>g</i>	Slightly blended	Very complete
<i>g#</i> or <i>ab</i>	Clearer than some	Self-complacent
<i>a</i>	Ringling quality	Powerful and brilliant
<i>a#</i> or <i>bb</i>	Blended	Wistful, appealing
<i>b</i>	Clear	Self-sufficient

Remarks

Effect different on different pianos

The following introspections were recorded by Observer G<sup>2</sup> after the experiment in identification of tuning forks: "These identifications were much harder than those of piano or violin notes. Attention was mainly on the pitch, but there was also a sort of vocal motor reaction aiding the pure memory for pitch. A comparison with imagined pitches a few vibrations higher or lower always took place in order to suggest or to verify the vibration rate. I always thought of the standard pitch (435 v.d.) as slightly flat. As I struck the fork I tried to observe any relaxing or tightening motor sensations, or any flatness or sharpness in comparing it with an auditory image. The four highest forks seemed distinctly high or sharp. It seemed almost impossible to distinguish between the first four or five forks whose differences are only one-half or one vibration."

<sup>2</sup>Observer G's threshold for pitch discriminations obtained by means of the Seashore Columbia Record disc was 98 per cent.



These introspective reports emphasize the fact that pitch is only one of many elements that enter into judgments of musical notes. Other important factors are the variations in apparent width and intensity of sound as you pass from the low tones to the higher ones. In some cases what seem to be directly ordinal judgments occur. The judgments are based upon a knowledge of general position and tonal spacing. Practice tended toward improving octave discriminations more than discrimination among the individual notes.

Frequent methods employed to determine the notes were, reference of all notes to an imagined  $c^1$  or  $a^1$ ; reference of all notes to their representatives in the middle octave; placing the note in a scale within the vocal range; placing the note in a chord followed by various harmonic progressions. Judgments were sometimes based on the so-called finished quality of sound which certain notes seem to possess.  $c^1$  is characterized in this way most frequently. Observers report "snap judgments," or immediate judgments, followed by an attempt at verification by indirect means. A few observers in Group 2 and more in Group 1 had feelings of familiarity for certain notes leading to immediate identifications as a result of their practice.

During the practice series more certainty of judgment was developed in observers who began the experiment with mere guesses. Direct judgments not infrequently occurred during the progress of the tests. A knowledge of octave differences came first in most cases, and later a feeling for individual notes.

Some observers attempted to build up a memory for separate notes by the formation of non-essential associations, such as: *B* has a funny sound when the key is lifted. In  $c\sharp$  the second overtone rings out.  $f\sharp$  rattles when played softly. Noticed where low notes begin to have distinct overtones and where high notes disappear into their more prominent overtones. The association that *c*, *e* and *g* have a finished quality may be the result of some habitual motor or perceptual reaction to these sounds. The notion that several observers had that certain notes are not quite in tune indicates also that they may have established a temporary or moderately lasting 'set,' auditory or motor, which gives rise to the feelings of the 'in-tuneness' or the flatness.

The introspections following the experiment with the



tuning forks emphasize again the effectiveness of a possible auditory or motor neural adjustment. This gives rise to an immediate judgment of higher, lower or the same, eliminating the comparison judgment with the aid of images.

The effect of the different distributions of black and white keys in the various chord combinations is illustrated by the observers' characterizations of the chords. The tonic chords built up on *c*, *f*, and *g*, which possess no accidentals are uniformly called hollow or complete. The rise in pitch also changes the feeling-effect for the chord. The higher pitches are thought of as more ringing or powerful.

After the lapse of a year with no deliberate effort to preserve any of the previously acquired ability to identify piano notes it is found that a number of essentials are still retained. The observers could still make use of their knowledge of octave differences, or more accurately differences in volume, for, as was remarked, some octaves have more than one change in them. In some cases the mental set for *c*'s remained, although other adjustments were much less clear. Former inaccuracies of the attention were recalled to aid this later test. Interest in this sort of tonal judgment was retained by one observer, who found herself constantly trying to tell the key of a piece and to pick out individual notes.

## SUMMARY AND CONCLUSIONS

a. *Effects of practice* Practice has a marked effect on an individual's ability to identify musical notes, that is distinct from the initial capacity of the individual, or from the type of instrument used in producing the notes. The observers in these tests improved irregularly from trial to trial in identifying piano notes, violin notes and the fundamental notes in tonic chords. A short period of practice covering an interval of about eight weeks with two ten minute practice periods a week brings an increase in the number of correct identifications and a corresponding decrease in the size of the errors made. Longer periods show continued improvement more or less irregular in its progress. Little sign of fixed plateaus is apparent in any of the results. The amount retained after the lapse of a year without practice is in some measure proportional to the degree of ability acquired through practice. In most cases some of the ability is lost over this period, although in a few cases progress rather than forgetting seems to have taken place.

b. *Group characteristics* Baird's results from tests on people possessing a high degree of ability to identify pitches show that "cases of absolute pitch memory may be represented as a series of gradations which extends from normality at the one extreme to a high degree of deviation from normality at the other extreme." The results presented on the preceding pages of this paper show this to be true not only within the limits of a selected group, but in a much wider field. The series of gradations extends without recognizable break below the obvious case of absolute pitch memory through all ranks of individuals who are familiar with our diatonic scale down to those who are tone-deaf.

The curve representing the frequency distribution for errors measured in semitones forms a fairly typical curve of normal distribution, and includes all grades of responses from a case in which 83 of the 88 piano notes were correctly identified down to cases in which only one tone is accurately named. The curve for the frequency distribution of correct notes is skewed a little toward the lower end and remains some-

what so after practice. To explain the absence of an occasional exceptionally bad case to offset the occasional very good cases the limitations of sampling within the group has been suggested. The error curve represents the individual's ability in a more comprehensive way and is less affected by random correct guesses.

The attempt to reduce the size of errors shows no tendency toward preferred intervals, such as those of a fourth or a fifth. On the contrary we find the overwhelming majority of the errors to be those of a semitone, and the larger the error the less its frequency in consecutive order from the halftone error to those over two octaves in magnitude. In a few cases octave errors show a slight increase in number over the errors of the major seventh, and in still fewer cases the rise in number of octave errors brings the frequency of this interval error next in order to that of the semitone or the whole tone.

The naturals on the piano all have a higher frequency of correctness than the accidentals. The notes at the extremes of the keyboard and those in the middle octave were most often named correctly. On the whole the accidentals in the thrice-accented octave received the fewest correct responses. The order of frequency of identifications for the seven octaves is as follows: four-accented, once-accented, contra, twice-accented, small, thrice-accented and great octave.

c. *Individual differences* The practice curves for the five observers from Group 1 show considerable variation in progress. With the observers who continued the experiment very nearly throughout the year, it holds true that the longer the period for practice, the smaller the error, irrespective of the initial error. The irregularities in the curves are due partly to changes in method of approach to the problem, partly to variations in attention and fatigue. Interest wavered at times, then received new spurts as the experiment progressed, and lagged again.

The time curves are more fluctuating than the error curves. These same factors affect the results here also. Accuracy, however, was made much more the goal in these tests than speed, so that reductions in the identification time are more or less incidental. The correlations between the average time per trial and the average error for each of these five observers are negligible. The correlation between the size of the error



per note and the time per response for one observer is zero. On the whole the time for incorrect responses is slightly longer than that for correct responses.

d. *Effect of timbre* The identification of notes from different instruments presents new situations which are more or less adequately met by the observers. The results from experiments on the pipe-organ using the four stops, viola, diapason, oboe and flute do not reveal striking differences between the average error or the average time per note for judging piano notes and for judging the notes of each of these stops. Since these tests were made after the observers had had several months of practice on the piano, it is possible that some of their methods gained through special practice with piano notes were transferred to the organ notes and aided identification. This may account for the fact that these observers show less variability than is usually the case in estimating notes of different timbre.

In the experiments judging the fundamental notes of various tonic chords, the observers found the task more difficult than identifying single notes. This contradicts the reports of a number of writers. The fusion of sounds with their new intensity orders confused the observers and made chord analysis necessary before the fundamental note could be identified. Small interval errors resulted rather than errors of the third or fifth as might have been expected.

The tests with tuning forks indicate some ability to remember the tone of relatively simple sounds over long periods of time. The forks with the larger vibration differences were better discriminated than those with small differences. These identifications seemed to be more dependent on the factor of memory for pitch than on recognition of tonal volume or brightness changes within these narrow limits. This pitch memory was expressed in both auditory and kinaesthetic terms.

e. *Kinaesthesia during the judging period* Good auditory imagery seems important in recognizing tones, more so, these experiments indicate than kinaesthesia. This fact is brought out by comparing the results from the tests for vividness of imagery with the individual's ranking in the tonal identification tests. It is further illustrated by the results from the kymograph records of movement of the larynx and



the tongue, and the breathing curve during the judgment periods. The usual shallow irregular breathing curve representing quick adjustments of the attention occurred, together with noticeable cessation of movements of the tongue and the larynx.

f. *Reproduction of notes* Ability to reproduce accurately by means of a tonoscope the vibration rate of designated notes depends partly on the ability to image the correct note and partly on one's technique in singing the imaged tone. On the whole in these tests the abilities to reproduce and to identify notes correctly run parallel. The experiments in singing the notes took place at the end of a year's practice in identifying notes. The earlier experiment had laid emphasis on auditory images and must have aided somewhat in this later experiment.

*Criteria for tonal judgments* Writers are agreed that frequency of hearing certain notes can not alone be made the basis for tonal memory. The results of special practice are held to be important by some writers and negligible by others. Many different theories have been advanced to explain the type of association involved in the identification of notes. Practice and memory plus an individual coefficient are made the basis for correct identification by Stumpf. A kind of limited association with particular pitch-blends is responsible according to von Kries for memory for pitch, because the instruments most frequently heard are not invariably the ones remembered. The ability is attributed by Abraham to associations of pitch built up in the auditory realm. In some people the sensory equipment may be finer so that finer differences of tone quality are recognized, or certain association paths in the brain may be more numerous or more closely knit together than in other people. The mechanism of the instrument producing the sounds may be such as to give to the different notes definite coloring which becomes associated with them, or the structure of the ear itself may affect the character of certain notes.

The tonal series may be regarded in two ways, first as a qualitative series, constituted by variations in pitch, and second as a quantitative series with the single quality of tone which may vary in ordinal terms, namely in pitch. The first interpretation is adopted by Boggs, who puts pitch on a par

with color in the visual series. One may be color blind and still recognize brightness differences, so one may be pitch deaf and recognize only the brightness differences in tones. Memory for absolute pitch is, therefore, not the result of learned associations but the result of a capacity for elementary experiences. Similarly Baird regards the recognition as dependent on the ability to observe in each note some distinctive characteristic c-ness, d-ness, etc., which can not be learned, but simply perceived if we are endowed with the capacity.

Other non-pitch attributes of tone have been stressed by writers, who have found experimentally that keen discrimination of tone-body or tone-color may with practice lead to a memory for the correct names of notes in the same way as pitch. The so-called attributes of tonality and vocality have an important bearing in memory for tones, for if attributive, they would greatly affect one's method of approach to the problem of absolute pitch memory. But it has yet to be proved that the similarity of octaves is attributive in that vowel sounds are not themselves the result of a previous association. The experimental evidence is strongly in favor of the latter view.

Watt advances the second interpretation of the tonal series as a quantitative system and regards pitch as an attribute of order. According to this, absolute ear is the natural condition of hearing, just as localization in any sense realm is a natural development from the primary attribute of order. This tendency in hearing is counteracted by the emphasis which music places on harmony and key relationships which are purely arbitrary and result according to Watt from the adoption of the volumic pattern of the octave as the standard of the whole range. People who are able to retain their original absoluteness in spite of the pressure toward relativity possess an unusual auditory 'disposition' or a greater delicacy of volumic outline and predominance.

The results from the present investigation support the view that certain factors other than pitch have importance in tonal judgments. The variations in the brightness and the volume of the notes afforded a cue that was used by all the observers. These factors were obvious to the poorest observer and were readily synthesized by the non-musical with the data from other sensory experiences. No evidence of aid from a recognition of vowel qualities occurs in these tests.



Reports of differences between octaves are always translatable in terms of brightness or of extensiveness, or of certain incidental features. Most of the observers reported a likeness existing between the same notes in the different octaves, at least in the octaves toward the middle range. At the extremes this similarity disappeared for some. Confusions in identification arising from the octave illusion are rare, only three observers showing high susceptibility to this error. Judgments of tonal position in a large number of cases involved naming a note within its octave in a single act, instead of dividing the process into first naming a note and then placing it into a particular octave, or vice versa, although all three methods were used to some extent during the course of the experiments. Observers occasionally confused volume differences with pitch differences, as for example when  $g^1$  was played, the observer knew the note was lower than  $g\sharp^1$ , but was uncertain whether lower was to be interpreted in volume and the note named  $g\sharp$  or in pitch and the note named  $g^1$ .

Complexity of sound was not found to aid in these experiments. Chords, in which no single predominance of pitch occurred, were more difficult to recognize than single notes. The identifications of notes from the four organ stops vary little from the results with the piano and the violin. Even when complexity of sound was entirely eliminated as in the test with the tuning forks, correct identifications and relatively small errors in judgment were possible.

The experiments bring out clearly the importance of frequency of hearing notes, provided attention is directed toward essential tonal characteristics. First, the individual notes assigned to the observers in Group 2 for special practice had a much higher frequency of correctness than the other notes; and second, practice over the entire keyboard produced a marked improvement in efficiency both as regards general reduction in error and increase in the number of correct responses. Furthermore, special practice correlates higher with improvement than the number of years of previous musical training.

The recognitive consciousness in identifying musical tones may be interpreted as similar to that in any typical judgment process. It is based on frequency, recency and satisfyingness of the individual's experience and is subject to the usual laws

of practice and forgetting. Interest plays an important part and is probably dependent on the delicacy of the sense organ and the wealth of clear auditory images. Without these factors attention tends to wander to extraneous aspects of the sound and learning will not proceed. If the observer is interested he will frequently attend to the correct factors without being analytically conscious of his task.

In acquiring a memory for absolute tone the observer does not remember well one or several notes and the others not at all, but he builds up a more or less cohesive structure about a few or many notes. Octave similarities, or the series of repeated relations within the octaves and the likenesses, existing between notes an octave apart, whatever their origin, make the task of apprehending the tonal series of eighty-eight steps much easier. These recurrent relationships may be wholly arbitrary, fixed in our minds by repetition, or they may be elementary. For many of the observers the tonal series did not contain very distinctive recurrences of chord or scale relationships from octave to octave. The series was one-dimensional without these regularly repeated units which facilitate the task of comprehending so long a range. Instead of calling the volumic series one of brightness changes, and the pitch series one of qualitative changes, it seems more appropriate to regard tones as belonging to a discrete series which the mind many apprehend in its diversity or as a unified whole consisting of even divisions of smaller units. By repeatedly directing the attention to the unity of the series, one may acquire the mental adjustment which finds little trouble in quickly assigning to random notes a position in a relatively well-organized whole.

This system of sounds when well incorporated is not a rigid inflexible affair. If an instrument is tuned lower or higher within limits (so that the tension of the strings is not altered too greatly), the system can be adjusted to the new situation and correct responses made. Also an observer may have given the correct name to a note in many previous successive trials, and at the next one place it a half tone higher or lower than it should be.

In immediate correct judgments of notes the observer has already established an accurate and (moderately lasting) 'set' or adjustment toward the tonal system so that responses then occur without reflection or comparison. It is conceivable that



the set may be the result of vocal motor reaction arcs aroused by the stimulus, or it may be wholly auditory. Whatever its pattern the feeling of familiarity, or the feeling of completeness or relaxation which accompanies the sound of certain notes, as  $c^1$ , can be explained as the result of definite establishment of associative connections, which with repetition will tend to drop out of consciousness. Thus one is not born with a sense for absolute pitch which is lost as a result of musical relativity but one learns to acquire a memory for individual notes through the analytic and synthetic processes involved in apprehending the originally discrete tonal series as an organization of recurrent parts and the development of well established associative connections giving rise to relatively permanent mental adjustments for the various musical tones.

## VITA

On August 8, 1892, I was born in Bayonne, New Jersey. In 1898 the family moved to Mount Vernon, New York, which has since been our home. I attended the public schools of that city and in 1909 entered Wellesley College by certificate. There I believe I was the first to take the special course covering five years with the study of the violin as my major interest. In 1914 I received a Certificate of Music and the A. B. degree. At that time I was offered a graduate assistantship for two years in the Department of Philosophy and Psychology at Wellesley to study under the direction of Miss M. W. Calkins and Miss E. A. McC. Gamble, and to assist the latter. The second year I went one day a week to Clark University at Worcester, Massachusetts, to take courses under John W. Baird and Dr. S. W. Fernberger. In 1916 I presented a thesis on the "Effects of Practice in Learning to Write New York Point for the Blind" and received the A. M. degree. The following year I spent at home, taking some courses in philosophy and psychology at Colombia University, and continuing my study of the violin. In 1917 when the need for men in war service arose, I was offered a temporary position as instructor in the Department of Philosophy and Psychology at Smith College. There I began the experimental work embodied in the present dissertation. The next summer I completed the requirement of academic points for the doctorate. My teachers at the University were Professors R. S. Woodworth, W. T. Bush, H. O. Ruger, and Dr. A. T. Poffenberger. In January, 1920, I was granted the degree of Doctor of Philosophy from Columbia University. I continued as instructor in psychology at Smith College until June, 1920, when I resigned to be married. The January, 1920, issue of the American Journal of Psychology and the April, 1921, issue contain accounts of experimental work done in co-operation with the students in the Smith College Laboratory on the "Tonoscope as a Means for Registering Combinational Tones."

EVELYN GOUGH BACON.

# AN EXPERIMENTAL STUDY OF SILENT THINKING

BY

RUTH S. CLARK, Ph. D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

No. 48

COLUMBIA UNIVERSITY CONTRIBUTIONS TO PHILOSOPHY AND PSYCHOLOGY

NEW YORK

APRIL, 1922

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Carey St., W. C.); Paris (16, rue  
de Condé)





## TABLE OF CONTENTS

	Page
1. Purpose of this investigation .....	5
2. Literature.	
A. Ach, N. Ueber die Willenstätigkeit und das Denken .....	5
B. Messer, A. Experimentelle-psychologische Untersuchungen ueber das Denken. Arch f.d. ges. Psych. 1906, VIII. pp. 1-224. ....	6
C. Bühler, K. Tatsachen und Probleme zu einer Psychologie des Denkvorgänge. Arch f.d. ges. Psych. 1907, IX. pp. 297-365 .....	7
D. Woodworth, R. S. Imageless Thought. J. of Phil. Psych. etc. 1906, III. pp. 701-708 .....	8
E. Störring, G. Experimentelle Untersuchungen ueber einfache Schlussprozesse. Arch. f.d. ges. Psych. 1908, IX. pp. 1-127 .....	8
F. Lunniss, B. A. A Study of Thought Processes. Studies from Bedford College, Univ. of London, 1914, pp. 67-161. ....	10
G. Betts, G. H. The Distribution and Function of Imagery, T. C. Contr. to Ed. No. 26, 1909 .....	11
H. Fernald, M. R. The Diagnosis of Mental Imagery. Psych. Rev. Mon. 1912. XIV. No. 58 ..	12
3. General Procedure .....	14
4. Results.	
A. Imagery Types .....	20
B. Sitting I. ....	27
C. Sitting II. ....	36
D. Sitting III. ....	46
E. Sitting IV. ....	56
F. Sitting V. ....	63
5. Individual Differences in Thinking .....	72
6. General Conclusions .....	84
Appendix .....	89
Bibliography of References .....	100





# AN EXPERIMENTAL STUDY OF SILENT THINKING

---

## CHAPTER I.

### INTRODUCTION

IN order to understand thinking from a point of view other than logical, it is necessary that thought processes be submitted to experimentation and be carefully observed in respect to their inherent characteristics. The results described below have been based upon an experimental study of thinking in relation to its physiological and psychological aspects. Not only has thinking in this investigation been submitted to observation under controlled conditions, but objective measures have also been made of its various characteristics, so that as little dependence as possible has been placed upon the subjective estimates of those thought processes which were being studied.

The term "silent thinking" has been used throughout this report to describe the processes carried on by the mind in its attempt to reach new conclusions, which are controlled in their development by the demands of a problematic situation. Subjects have been asked to solve problems without consciously making use of any explicit form of behavior. The value of the solution was not taken into account. Any solution offered by the subject, which to him was new, was accepted. It is such processes as these that have been submitted to experimentation in this investigation and that have been studied from the psychological and physiological view points.

There have been few attempts to deliberately observe under controlled conditions the silent reaching of new conclusions. Ach, in his study of thought,<sup>1</sup> was interested to discover whether extra-conscious determining tendencies, set up by a given task, could yield new products which would not be the result of reproduction. He exposed to the view of

---

<sup>1</sup>Ach, N. *Über die Willenstätigkeit und das Denken*, 1905, p. 196.

his subjects separate nonsense syllables to which the subjects were at times required to give rhymes, at times examples of alliteration. For each experiment he recorded the time and retrospections of the subject. As a result, he stated that new conclusions could be reached by the influence of extra-conscious determining tendencies, and that the novelty of the conclusions might be due to successive association, as when the given stimulus was followed by the response for the first time, or to simultaneous association, as when a syllable made up of a new combination of letters was produced by the subject.<sup>1</sup>

Messer, in his experimental investigation of thinking,<sup>2</sup> stated that judgments may differ in relation to other conscious processes in the degree of novelty which they possess.<sup>3</sup> Some of the tests in certain series of his experiments produced conclusions which were characterized as novel by the subjects. From a study of these reactions, he found "activity" especially characteristic of original thinking.<sup>4</sup> This activity he described as the voluntary production of relations and the mental processes of seeking, reflecting, proving, resolving, accepting and rejecting. The more original and the more prolonged the thinking, the more pronounced were these activities. When the thinking was especially hurried, as in passing rapidly over a minor premise to reach the conclusion, these mental processes were rarely reported. Later on in his report<sup>5</sup> Messer classified two of these activities, seeking and reflecting, as "*Bewusstseinslagen*,"<sup>6</sup> being imageless states. Whether or not the other activities were similar states, it is impossible to say. Messer also found<sup>7</sup> that the influence of the "*Aufgabe*" was more marked in original thinking than in reproduction. Something was "meant" and "intended" in the formation of original judgments which was not so characteristic of the recall of familiar material.

Another contribution from the Würzburg school was that

<sup>1</sup>Ib. p. 208.

<sup>2</sup>Messer, A. Experimentelle-psychologische Untersuchungen über das Denken. Arch. f. d. ges. Psych. 1916 VIII. pp. 1-224.

<sup>3</sup>Ib. p. 115.

<sup>4</sup>Ib. p. 125.

<sup>5</sup>Ib. p. 183.

<sup>6</sup>This term, suggested by Marbe, appeared for the first time in the work of his pupils, Mayer A. and Orth, J., Zur qualitativen Untersuchungen der Association. Zsch. f. Psych. 1901. XXVI. pp. 1-13.

<sup>7</sup>Ib. p. 126.

of Bühler.<sup>1</sup> He purposefully limited his observations of thinking to the study of the reaching of new conclusions, eliminating reactions which were mechanical or even familiar. In order to arouse in his subjects original thought, he read to them passages from Nietzsche, Marie vonEbner, Sachenbach, Pauly and Otto Weiss, requesting them to take some attitude toward them, answering either yes or no. He then recorded their reaction times and retrospections.

Bühler found, as did Ach and Messer, that thinking was dependent upon the influence of determining tendencies set up by the task.<sup>2</sup> In respect to the constituent parts of the thought processes, the subjects reported the presence of sensory images, both verbal and concrete. Inner speech, however, did not invariably appear.<sup>3</sup> The subjects also described sensory experiences of spatial directions and of changes in direction.<sup>4</sup> "Bewusstseinslagen" were reported as well, such as feelings of doubt, astonishment, and expectancy. But these were not all. There were certain states of mind, "Gedanken," "thoughts about" this or that object, which were described as vivid, convincing, and definite, although they contained no sensory elements. Ach had found similar states of mind reported by his subjects.<sup>5</sup> These he had termed "Bewusstheiten." Such states of mind were also reported by Messer.<sup>6</sup> He described them as "unanschaulichen Wissen" and included them under the head of "Bewusstseinslagen."

Imagery and "thoughts" Bühler decided were the main factors which guided thinking.<sup>7</sup> Imagery, however, was found to have been too fragmentary and sporadic to be the basis of continuous thinking. Since thinking was as distinct with images as without, and in many instances was remembered, while the images which accompanied it could not be recalled, he concluded that "thoughts about" objects were the fundamental constituents of the thought process. He believed that these thoughts were irreducible units, resisting analysis. Three types of these thoughts were recognizable: the consciousness of a universal principle, the consciousness of

<sup>1</sup>Bühler, K. *Tatsachen und Probleme zu einer Psychologie des Denk-  
vorgänge*. Arch. f. d. ges. Psych. 1907. LX. pp. 297-365.

<sup>2</sup>Ib. p. 315.

<sup>3</sup>Ib. p. 319.

<sup>4</sup>Ib. p. 315.

<sup>5</sup>Op. cit. p. 210.

<sup>6</sup>Op. cit. p. 175.

<sup>7</sup>Op. cit. p. 316.



the relation between one part of thought and another, and the consciousness of the relation of the thinker to the thought, "intention." The difficulty lay in his inability to describe these various kinds of thought. Bühler believed that the "thought about" something sensory or imaginal was neither sensory nor imaginal, as an object and its meaning are distinct. He therefore concluded that "knowledge about" something involved an added problem: the description of that knowledge, which was still to be determined.

Similar results were simultaneously obtained in this country from the experimental work carried on by Woodworth.<sup>1</sup> His subjects were asked questions which were considered difficult enough to stimulate real thinking. Ordinarily the subjects were asked to give their retrospections at the completion of their solutions, but frequently the experimenter interrupted them before they had reached their conclusions, asking: "Any visual pictures?" "Any words heard?" "Any feelings of bodily movement?" From the study of these results and from his own introspections, Woodworth concluded that thinking contains elements "wholly irreducible to sensory terms." Frequently the content of thinking was merely a "particular feeling of the thought in question," a state of mind apparently identical with Bühler's "*Gedanke*." In describing one of his own thought processes, Woodworth was even able to add: "When thought is really effective, when some new insight is gained, the words are absent, though they soon come tagging after."

An experimental investigation of inference was carried out in Zurich by Störriing.<sup>2</sup> When the simple syllogisms were presented, Störriing found that, after perceiving the propositions, the subject quickly received an idea of the whole problem by means of visual imagery, the middle term being recognized as common to both premises. The conclusion seemed to follow immediately, as if it were a direct result of the perception.<sup>3</sup>

In syllogisms where more complex relations were introduced, however, the drawing of the conclusion was dependent upon the mediation of other processes, by means of which

<sup>1</sup>Woodworth, R. S. *Imageless Thought*. Journ. of Phil. Psych. etc. 1906, 111. pp. 701-708.

<sup>2</sup>Störriing, G. *Experimentelle Untersuchungen über einfache Schlussprozesse*. Arch. f. d. ges. Psych. 1908, XL. pp. 1-127.

<sup>3</sup>Ib. pp. 14 ff.



the relations of the middle term had to be made clear before the conclusion could be drawn. Occasionally the subjects reported a "synthesis of relationships" which presented a view of the whole so clearly that the conclusion followed directly.<sup>1</sup> In some cases the subjects reported the presence of a "Gedanke" of the progress of the thought,<sup>2</sup> before drawing their conclusions.

Other contents of consciousness described by the subjects included feelings; images, both concrete and verbal; sensations of movement in the eyes, hands, and chest; "Bewusstseins" of certainty, identity, possibility, opposition, etc.; while the "Gedanken" ranged from simple thoughts of relations to the consciousness of universality. The term "Bewusstseinslage" was not used by the subjects in their reports or by the author in classifying their imageless states. The terms "Bewusstsein" and "Gedanke" were preferred, and at times use interchangeably.<sup>3</sup> An attempt, moreover, was made to study further the nature of the *Gedanke*. The subject was asked not to react until his thoughts about relations were as clear in his mind as possible. As a result, concrete and verbal imagery and sensations of movement were frequently reported.<sup>4</sup> Other results, however, showed that when similar syllogisms were repeatedly presented to the same subject, in some cases the sensory elements disappeared, while in other cases they became more prominent.<sup>5</sup> When the subjects were asked to react as quickly as possible, the imagery was usually reported less clear.<sup>6</sup>

Imagery apparently played an important part in the thinking of his subjects, for they frequently reported that they used images of lines, circles, and cubes in order to make the meaning of the relationship between the terms more clear.<sup>7</sup> However, imagery could not have been necessary to the development of the conclusion, for in some instances it was absent,<sup>8</sup> while at other times it was too weak to produce a

---

<sup>1</sup>Ib. p. 22.

<sup>2</sup>Ib. p. 27.

<sup>3</sup>Ib. pp. 26, 27.

<sup>4</sup>Ib. pp. 50, 69, 73, 112.

<sup>5</sup>Ib. pp. 76, 105.

<sup>6</sup>Ib. pp. 17 ff.

<sup>7</sup>Ib. pp. 10, 53, 81.

<sup>8</sup>Ib. pp. 9, 28.

conclusion, even if a feeling of certainty accompanied the conclusion reached.<sup>1</sup>

Finally, in Störing's results there were more references made to the part played by movement than in any of the works of his predecessors. Eye movements were reported, not only in perceiving the positions of the various terms, but in connection with visual imagery.<sup>2</sup> Arm movements seemed to represent relationships of left and right.<sup>3</sup> The movements caused by breathing were also noted, but not measured.<sup>4</sup> Movements, he concluded, either disappeared in the presence of imagery, or they were overlooked in the description because of the attention given to the image.<sup>5</sup>

An investigation carried on by Miss Lunniss at Bedford<sup>6</sup> closely resembles that of Bühler's. Short problems were given the subjects to which each replied either affirmatively or negatively. The time for each reaction was taken and the subjects recorded all that had taken place in consciousness from the time of hearing the question to the time of giving the answer. The experimenter then gained what further information was necessary to make these protocols clear by asking "judicious questions," and recording the answers.

Imageless thought was found to have characterized 26 of the 123 reactions, the remaining responses having included imagery.<sup>7</sup> Imagery was found to have been necessary to original thinking for one subject, to have been occasionally necessary in the case of another subject, and quite irrelevant to thinking for the third subject. No connection could be found between the time expended in thinking and imagery, for short reactions were as apt to have imagery as not.<sup>8</sup> In the case of two subjects, difficulty in the thinking seemed to account for imagery in some cases.<sup>9</sup>

Instances of Bühler's three kinds of "Gedanke" were also found.<sup>10</sup> These mental processes were not described as such by the subjects, but were interpreted in this way by the ex-

<sup>1</sup>Ib. p. 6.

<sup>2</sup>Ib. pp. 24, 25, 42, 107, 125.

<sup>3</sup>Ib. pp. 17, 42, 44, 45.

<sup>4</sup>Ib. p. 56.

<sup>5</sup>Ib. pp. 29, 59.

<sup>6</sup>Lunniss, B. A. *A Study of Thought Processes*. Psych. Studies from Bedford College, Univ. of London, 1914. pp. 67-161.

<sup>7</sup>Ib. p. 76.

<sup>8</sup>Ib. p. 80.

<sup>9</sup>Ib. p. 81.

<sup>10</sup>Ib. pp. 85 ff.

perimenter after an analysis of the subjects' protocols. They seemed to be an essential part of the judgment process. Moreover, there were no "ways of knowing" which played a part in the thinking which could not be included within one of the three classes described by Bühler, provided they were interpreted broadly enough by the experimenter.<sup>1</sup>

Various "Bewusstseinslagen" were also described by the subjects.<sup>2</sup> In fact, the experimenter found that an attempt to enumerate them all would be too confusing. In addition to those already described by Messer<sup>3</sup> and Bühler,<sup>4</sup> the experimenter found a distinct "consciousness of the manner of mental procedure," a "questioning attitude" and a "laboratory attitude." These "conscious attitudes" seemed to have been distinct from the "ways of knowing" in the minds of the subjects, and yet the experimenter acknowledged that the two kinds shaded off very closely into one another.<sup>5</sup> Unlike the "ways of knowing," however, the "conscious attitudes" could be absent when judgments were found and so were obviously not essential to the thinking.

Finally, the evidence which more conclusively showed the value of Bühler's "ways of knowing" was that reactions for which no formation of judgments were required were characterized by the lack of these "ways of knowing," especially the "consciousness of the relation of the present data offered for apprehension to some whole already apprehended."<sup>6</sup>

Betts, in an experimental study of imagery,<sup>7</sup> included an investigation of the part played by images in thinking. After having given his subjects a problem to solve, he asked them to write, instead of a description of their solution, an account of the mental processes which had meanwhile been taking place. The subjects were skilled in introspection and were used to this method of interrupted thinking. The following mental processes were described: imagery, sensory elements, personal feelings, elements of relation and of meaning, feelings of effort, strain and attention.<sup>8</sup>

<sup>1</sup>Ib. p. 100.

<sup>2</sup>Ib. pp. 117 ff.

<sup>3</sup>Op. cit. pp. 175 ff.

<sup>4</sup>Op. cit. p. 315.

<sup>5</sup>Op. cit. p. 123.

<sup>6</sup>Ib. p. 139.

<sup>7</sup>Betts, G. H. *The Distribution and Function of Imagery*. T. C. Contr. to Ed. No. 26, 1909.

<sup>8</sup>Ib. p. 62.



Imagery was not reported by all the subjects and by its very nature it seemed to be merely incidental to the thinking. The experimenter finally concluded that imagery had been overestimated as an important mental content.<sup>1</sup> Images seemed to appear only when thinking was "baffled" or when percepts would ordinarily have been of use.<sup>2</sup> Consequently the most "successful" thinking had apparently the least imagery.<sup>3</sup> The imageless states of consciousness were not analyzed by Betts.

Dr. Fernald, in her investigation of individual differences in imagery<sup>4</sup> included a study of the part played by imagery in the solving of problems. Problems in both mathematics and physics were given to the subjects.<sup>5</sup> This kind of problem was selected with the hope that concrete rather than verbal imagery would be aroused. The detailed results of these experiments were not published, but the experimenter's conclusions were that the individual's type of imagery, which was found to have characterized other kinds of mental processes, was found to prevail in thinking as well. Visual imagery, however, seemed more important to the solving of the mathematical problems, while verbal imagery was emphasized more in the solution of the physics problems.

Dr. Fernald also concluded that, in solving the most difficult problems, more attention was required for the performance of the work itself, while less attention could be paid to imagery and consequently a sparsity of images was reported.<sup>6</sup> This seemed a probable explanation of some imageless states. However, the presence of imagery was more frequently reported with an increase in the novelty and difficulty of the problem. In fact, the experimenter was able to devise problems which invariably required the presence of imagery by making the problem sufficiently difficult and at the same time denying perceptual aids to the subject. Therefore, imagery played an important part in original thought. But the experimenter did not deny the possibility of other factors being equally important or even more so than imagery.

From this survey of the studies which have been made of silent thinking, it is clear that very little has been done to ob-

<sup>1</sup>Ib. p. 98.

<sup>2</sup>Ib. p. 70.

<sup>3</sup>Ib. p. 94.

<sup>4</sup>Fernald, M. R. *The Diagnosis of Mental Imagery*, *Psych. Rev. Mon.* 1912. No. 58.

<sup>5</sup>Ib. p. 95 ff.

<sup>6</sup>Ib. pp. 36 ff.



serve the detailed characteristics of thought under experimental conditions. Problems have been given to subjects skilled in introspection, the time of the reaction has been recorded and the retrospections of the subject have been studied. The mental processes described by the subjects have been interpreted and classified by the experimenter and the relative importance of each to thinking has been determined. In general, mental images have been found to be unnecessary to the reaching of a novel conclusion. They often appear in thought processes, but they do not always seem relevant to the content of the thought. Other mental processes described by the subjects have been imageless states, most frequently characterized as conscious "attitudes" or as "thoughts about" some concrete object. These have been considered by most experimenters to be essential to thought. References to sensory states of consciousness during thinking, such as the awareness of movements, have been merely incidental to the reports of these experimenters and no attempt seems to have been made to study them or to note their importance.

## CHAPTER II.

### GENERAL PROCEDURE

In distinction to the above methods of procedure, the present investigation<sup>1</sup> has confined itself to a compilation of as many objective measures of the physiological and psychological aspects of thinking as possible, leaving to the experimenter little or no chance for rendering an interpretation of the results. As the bone of contention between the investigators has been based upon a consideration of the prevalence and importance of imageless states as compared with imagery, it was believed best in this investigation to use every means to arouse imagery, and to record the results.

Problems were given which were likely to require for their solution images<sup>2</sup> in the various sense fields, and the subjects were asked to give an account of the mental processes which occurred between the time of the presentation of the problem and the moment at which the subject was aware of an original conclusion. Their attention was especially called, in some sittings, to the presence of imagery during the solution of the problem. Only two subjects were conversant with the work from the Würzburg Laboratory in detail, while the other subjects, as their introspection showed, gave little evidence that they were familiar with imageless thought as such.

The experiment was carried on in a dark room lighted by but one bulb in the center, all other distracting objects having been removed from sight. A continuously recording kymograph was used which was run by electricity. The time in fifths of a second was given by means of a vibrating reed, and recorded by one of the needles. A second needle was connected with a key on the arm of the subject's chair while the third needle recorded the taps of the experimenter's key.

The apparatus was placed on a large table in the center of the room (See Fig. 1.) The motor and gear wheels were covered by a box lined with cotton in order to reduce their sound. All apparatus was screened from the view of the subject who sat about four feet away from the end of the table.

---

<sup>1</sup>This study was completed in 1919.

<sup>2</sup>Betts pointed out the prevalence of imagery in the solution of certain problems where percepts would ordinarily be of service. *The Distribution and Function of Imagery*, p. 70.

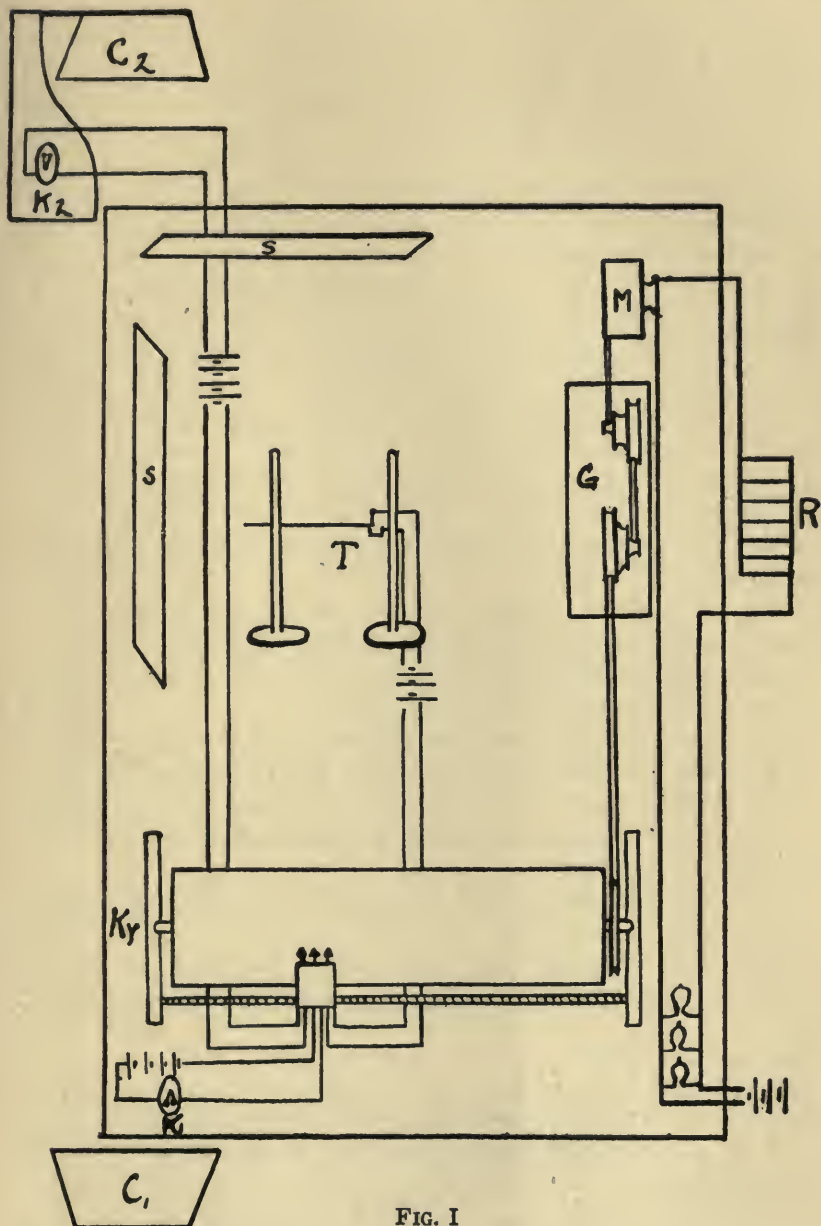


FIG. I

Ky. Kymograph.  
 R. Rheostat.  
 G. Gear Wheels.  
 M. Motor.  
 S. Screens.

C<sub>1</sub> Experimenter's Chair.  
 C<sub>2</sub> Subject's Chair.  
 K<sub>1</sub> Experimenter's key.  
 K<sub>2</sub> Subject's key.  
 T. Vibrating Reed.



The experimenter sat at the opposite end of the table facing the apparatus with her hand on the key, which was hidden from sight of the subject. Thus it was possible for the experimenter, who was facing the subject, to observe his movements and to indicate the nature of them on the kymographic record by pressing her key, making use of a code of signals previously determined.<sup>1</sup> As this key was out of sight of the subject, this description of movements on the part of the experimenter was carried on unknown to the subject, who believed that merely the time of the reaction was being recorded. The noise of the apparatus was not loud enough to interfere with the giving of the problem by the experimenter or the description of retrospections by the subject; it being a low rumble which seemed to offer no distraction. The subjects were ten graduate students in psychology at Columbia University; five men and five women.<sup>2</sup>

The problems were carefully selected to meet the following requirements:<sup>3</sup>

(1) They must demand no form of explicit behavior for their solution, silent thinking alone being desired.

(2) They must present a problem for which a novel conclusion could be devised by the subject.<sup>4</sup>

(3) They must not call for mere book knowledge where verbal elements would naturally play the important part.

(4) They must not be arithmetic problems, as fixed habits have usually grown up about this kind of mental activity and its products can hardly be termed novel.

(5) They must be stated simply so that no confusion can arise as to their meaning.

(6) They must not call merely for the answer yes or no, as the greatest possible variety in response was desired.

---

<sup>1</sup>See Appendix p. 89.

<sup>2</sup>To these subjects: The Misses Cornell, Corwin, Gough, Wilkins and Mrs. Berliner; the Messrs. Achilles, Brimhall, Gates, Jackson and May, I wish to express my deepest gratitude for their generous cooperation. I wish also to thank the Neurological Department of the School of Physicians and Surgeons for the use of their instruments; and especially their assistant, Mr. De Bra, for his help in recording the minute bodily movements. To Professors Woodworth and Cattell and to Dr. Poffenberger, I am likewise indebted for inspiration and helpful suggestions.

<sup>3</sup>See Appendix, p. 90 ff.

<sup>4</sup>If the problem was one familiar to the subject, it was eliminated. Moreover, as soon as the subject had reached his conclusion, he was asked if it were new to him. If not, the results were not included in this report.

(7) They must present situations, which, though not common, might happen in the experience of anyone, and so disclose mental processes which are ordinarily used to solve the problems of everyday life.

(8) They must not arouse emotional states any more than necessary as this would needlessly introduce extraneous mental processes.

As soon as the problem was given to the subject, objective records were taken on the kymograph during the thinking. After the subject had signalled to the experimenter that he had solved the problem, he was asked the following questions, always in the same order:

1. Will you describe the course of events?
2. Was the conclusion novel? How much so?
3. Was imagery present? Kind?
4. Was there more than one kind of imagery present at once? Describe.
5. Was imagery present throughout the thinking?
6. Was there a feeling of progress throughout? If not, Where?
7. Did your mind wander? When?
8. In what terms did the novel part of the conclusion present itself?
9. Did you at any time have a desire to speak?

In spite of the fact that many of these questions called for information which had already been given in answer to the first question, they were nevertheless always asked after each reaction. No other questions were asked and the answers to these were recorded verbatim. The subjects used their own terms to describe the contents of their thought processes. They were at no time influenced by suggestion to change these expressions or to search for any special kind of mental elements, with the exception of imagery, as noted above.

An attempt was first made to determine the individual differences in mental images of the various subjects. Each one was asked to respond to the Seashore Questionnaire,<sup>1</sup> giving a numerical estimate of the degree of vividness of each image. From these results a curve of each subject's capacity for vividness of each kind of imagery was drawn. The subject's type

<sup>1</sup>Seashore, C. E. *Elementary Experiments in Psychology* pp. 110 ff.



of verbal imagery was determined by his responses to pictures of familiar objects. The proportion of verbal imagery to concrete imagery was determined by the relative frequency of each to a list of twenty words, presented orally, which were later to appear in some of the problems.

Sitting I, during which problems were given, was not held until at least a week had expired, in order that specific recollections of the results of the imagery tests would not intrude themselves when similar or identical stimulus words appeared in the presentation of the problems. The kymographic record in this sitting included (a) the time of the thinking in fifths of a second and (b) a record of the movements made by the subject during the thinking. The description of these was made in code by the experimenter on the kymograph key. Five problems were solved by each subject in each sitting.

Sitting II was similar to Sitting I with the exception that the subject was given the added instructions to press the key on the arm of his chair whenever he was conscious of the presence of imagery, releasing it when the imagery disappeared. The kymographic record, therefore, included (a) the time of the thinking in fifths of a second, (b) a description of the subject's movements made by the experimenter, and (c) the presence and absence of imagery occurring throughout, as recorded by the subject.

In Sitting III the attempt was made to determine more exactly the amount of movement made by the subject. The experimenter, therefore, pressed her key whenever she observed any movement made by the subject, releasing it when the movement disappeared. The subject was asked to react in the same way as in Sitting II. The kymographic record, therefore, showed the simultaneous record of (a) the time in fifths of a second, (b) the presence and absence of movement, and (c) the presence and absence of imagery.

Throughout Sitting IV various distractions; visual, auditory, verbal and hand-motor, were introduced during the thinking in order to learn, if possible, the connection existing between imagery and movement. Each kind of distraction was introduced during the solving of a problem which was likely to require imagery of the same sense field. Kymographic records were made by the experimenter of the nature and degree of the effect produced by the distraction.

Finally, Sitting V was given in the laboratory of the Neuro-



logical Department of the School of Physicians and Surgeons. The subject was asked to solve certain problems while records of his respiration, both thoracic and abdominal, the volume changes of his arm, the horizontal and vertical movements of his larynx and the movements of his tongue were being simultaneously made upon the kymograph. Four subjects only were used during this sitting, twelve problems in all being solved.

## CHAPTER III

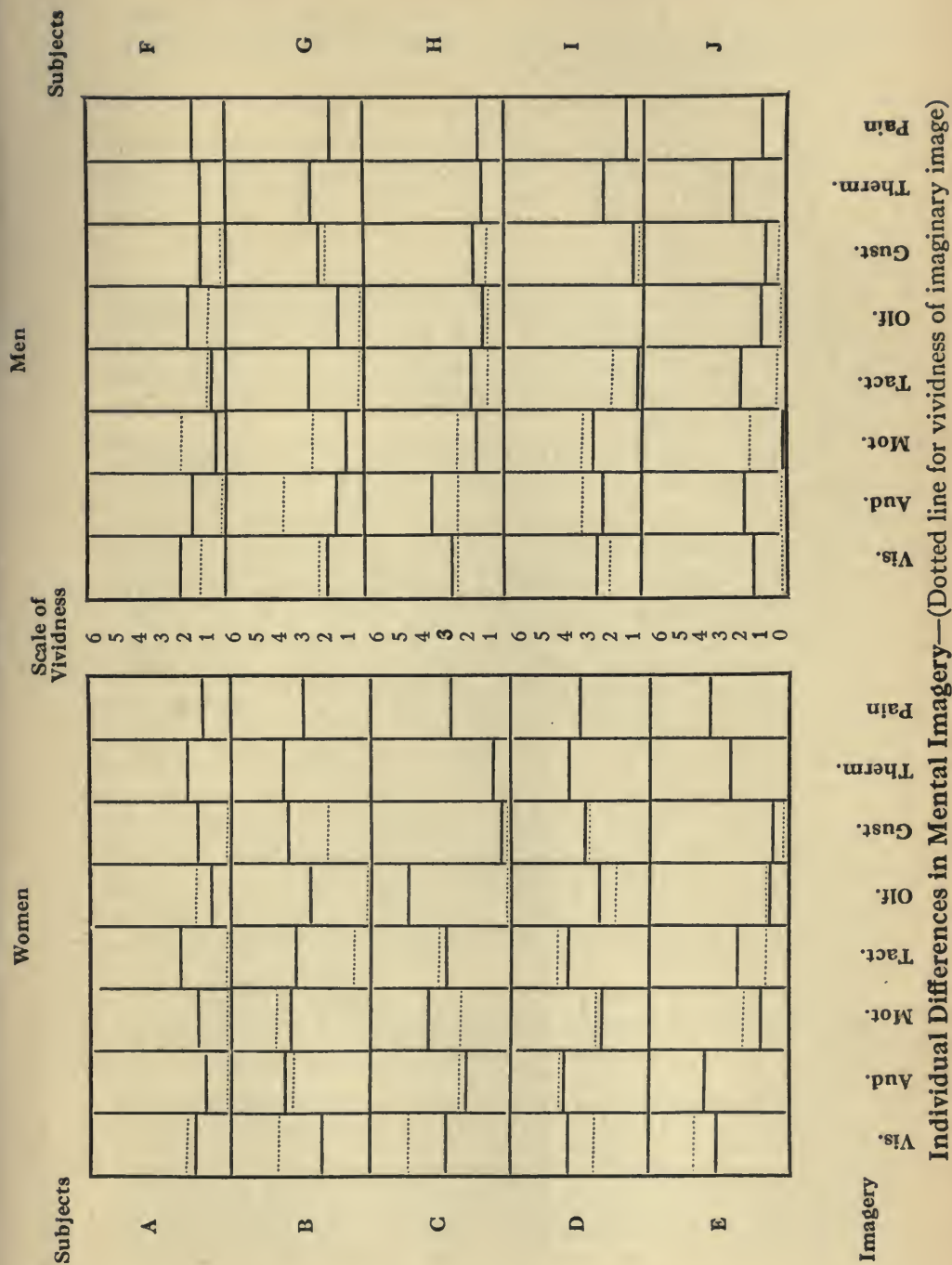
### 1.

#### INDIVIDUAL DIFFERENCES IN IMAGERY

Test 1. Seashore's original questionnaire was somewhat modified to secure certain desired results. It was made shorter by omitting all but three specific examples for each topic under each kind of imagery. The scale of five degrees of vividness was increased to six in order to eliminate a middle class into which the majority of images could be grouped. Moreover, for each kind of imagery, with the exception of those of temperature and pain, the subject was asked to "construct" an image, such as: a visual image of a bouquet of nasturtiums and pansies in a brass bowl, an auditory image of water dropping onto a piece of velvet, etc. A numerical estimate of the degree of vividness of each of these imaginary images was required, as in the case of the other images called for in the questionnaire.

The following diagrams, Fig. II, show the average degree of vividness with which each subject was able to summon each kind of imagery. It is apparent from these results that every subject could voluntarily call to mind at least one image of each sense field, with the sole exception of Subject I who was unable to summon any olfactory image. With this exception the range of imagery for all subjects included all the sense fields. In spite of this ability, a study of the separate reactions shows that Subjects D and H were the only ones able to summon every image called for in the questionnaire. The other subjects could not always call to mind images of familiar sensory experiences. One subject, J, was unable to summon twenty-seven such images. This inability to revive familiar sensory experiences may be due to the nature of the stimulus which was in each case visual verbal, the name of the sense experience. In each instance, however, the subject was "thinking about" the experience, realizing it to be familiar, without being able to reinstate it, though he could reinstate others of the same sense field.

A glance at these curves also shows a noticeable sex difference, the women reporting more vivid imagery of every kind





than the men.<sup>1</sup> The difference is made clear in the following table, Table I, where the average degree of vividness of each kind of imagery is compared.

TABLE I

Sex		Vis.	Aud.	Mot.	Tact.	Olf.	Gust.	Therm.	Pain
M.	Av.	1.7	1.5	.9	1.3	.8	.9	1.4	1.3
	A. D.	.3	.5	.5	.6	.4	.5	.4	.4
F.	Av.	2.5	2.5	2.	2.8	1.3	1.5	2.2	2.6
	A. D.	.8	1.2	1.2	.6	.9	1.2	1.2	.9

In spite of the large average deviations for many of the scores, there is very little overlapping in the degree of vividness in imagery for the sexes, as only one woman, Subject A, had the majority of her averages as low as those for the men, while no man had a majority of averages as high as those of the women.

All the subjects could construct at least one imaginary image. Eight constructed at least four. The gustatory imaginary image, (to imagine the taste of chicken broth flavored with mint) seemed to offer the greatest difficulty, only four subjects being able to construct it. One of these gustatory images, however, was described by the subject as "clear and vivid rather than vague and dim." Indeed every subject summoned at least one imaginary image which was more vivid than his average for that kind of image. In fact, the majority of the imaginary images, twenty-three of the forty described, were more vivid than the average of that kind of image for that subject.

Every subject later reported the presence of at least three imaginary images in the thinking out of his problems. Seventy-eight such images were described, sixteen of which were purely visual, the rest combining various kinds of images. One subject, Subject D, described as many as nineteen of these imaginary images.

Test II. The individual's type of imagery was taken in this investigation to mean the kind of imagery most frequently used by that individual.<sup>2</sup> Assuming that frequency in use

<sup>1</sup>Galton, F. *Inquiries into Human Faculty*, p. 99. Armstrong, A. C. and Judd, C. H. *The Imagery of American Students*, *Psych. Rev.* 1. pp. 496-505. French F. C. *Mental Imagery in Students*, *Psych. Rev.* IX. pp. 40-56. These suggest similar sex differences.

<sup>2</sup>Seashore, C. E. *Elementary Experiments in Psychology*, p. 114, and Feuchtwanger, A. *Versuche ueber Vorstellungs-Typen*. *Zsch. f. Psych.* 1911, LVIII. p. 162, emphasized this criterion for determining type of imagery.

implied ease and readiness in association, a list of twenty words<sup>1</sup> was read to the subject who was requested to write down the kind of imagery which first came to mind when each word was perceived. The results of this are given in Table II.

From the results in column I it is clear that in response to an auditory verbal stimulus most of the subjects showed themselves to be of the visual type.

Test III. In order to determine whether the type of imagery depended upon the nature of the stimulus, a second set of twenty stimuli, cards upon which were pasted outline sketches of familiar objects, was presented to each subject. These results are given in column 2, Table II. With these visual concrete stimuli, verbal elements came first to mind in the case of every subject, so that the type for each was verbal. From these results it seems apparent that the type of imagery possessed by an individual depends upon the nature of the stimulus.<sup>2</sup>

If this conclusion be true, the types determined by Test II should coincide with the "actual" type of each subject: that kind of imagery used most frequently by each subject in the solving of his problems, which were themselves responses made to auditory verbal stimuli. These results are also shown in Table II. Visual images predominated in the solving of the majority of problems for all subjects. Consequently those subjects which were found to have been visual in type by Test II remained so. Subjects B, D and F, however, used verbal imagery in thinking as frequently as they resorted to visual imagery. But Subjects G and H, who were verbal in type by Test II, showed themselves to be of a visual type in actual thinking. A test such as Test II cannot, therefore, adequately determine the type of imagery which an individual will actually use in thinking, even when responding to the same kind of a stimulus.

The twenty stimulus words of Test II were words which appeared later in the statement of the problems, thus they were presented again orally to the subject as stimuli. These words appeared in the statement of eighty-one problems

<sup>1</sup>See Appendix p. 95.

<sup>2</sup>Similar results were noted by Colvin, S. S. *Methods of Determining Ideational Types*, Psych. Bull. 1909, VI. p. 235; by Kakise, W. A. *Preliminary Experimental Study of the Conscious Concomitants of Understanding*, Am. J. of Psych. 1911, XXII. p. 63, and by Fernald, M. R. *A Diagnosis of Mental Imagery*, Psych. Rev. Mon.; 1912, XIV. No. 58. p. 98.

TABLE II

Subj.	Test II	Test III	"Actual" Type	"Supposed" Type
A	Vis.	Verb. (mot.)	Vis.	Vis.
B	Verb. (aud. mot.)	Verb. (aud. mot.)	Vis. & Verb. (aud. mot.)	Verb. (aud. mot.)
C	Vis.	Verb. (aud. mot.)	Vis.	Verb. (aud. mot.)
D	Vis.	Verb. (aud. mot.)	Vis. & Verb. (aud. mot.)	Vis.
E	Vis.	Verb. (aud. mot.)	Vis.	Vis.
F	Vis.	Verb. (aud.) (mot.)	Vis. & Verb. (aud. mot.)	Vis.
G	Verb. (aud.)	Verb. (aud.)	Vis.	Verb. (vis.)
H	Verb. (aud. vis.)	Verb. (aud. mot. vis.)	Vis.	Aud.
I	Vis.	Verb. (aud. mot.)	Vis.	Mot.
J	Vis.	Verb. (aud. vis.)	Vis.	Aud.



which were solved by the various subjects. In twenty-three of these instances the subjects did not report any imagery called to mind by these particular words. In fifty-eight instances, however, some imagery was described, of which thirty-three images, or 56.8 per cent, were the same kind of images as those aroused by the same stimulus word in Test II. It is possible, therefore, that a test such as Test II may indicate the kind of imagery which will immediately succeed a certain specific stimulus, but that it will not necessarily indicate the type of imagery which will predominate in an extended mental process.

After having taken these three tests, each subject was asked to state what he supposed to be his type of imagery. The responses to this inquiry are also given in Table II. Only half of the subjects showed in their thinking the dominant kind of imagery which they supposed their type to be. Again, only half of the subjects gave as their supposed type that which was disclosed by the results of Test II., where auditory verbal stimuli were presented. Moreover, only two subjects supposed their type to be that which, according to Test III, was their customary association to visual concrete stimuli. Finally, degree of vividness of imagery could not have been the basis of the subject's judgment of his supposed type, for only in the cases of Subjects F, H, and I was there correspondence between the "supposed" type and maximum vividness in imagery. An estimate of one's own type of imagery is, therefore, not necessarily influenced by results gained from an analysis of one's own imagery by means of tests, nor does it necessarily anticipate the kind of imagery which will predominate in the solving of problems.

From the above evidence, slight though it be, the following conclusions seem justifiable:

1. An individual may think about a familiar sensory experience and still not be able to voluntarily call to mind an image of it, although he may be able to summon other images in the same sense field.

2. Mental images for men are less vivid than for women.

3. Imaginary images can be voluntarily constructed and they are apt to be more vivid than the individual's average image in the same sense field.

4. The kind of imagery which first comes to mind varies with the kind of stimulus which arouses it and with the

specific characteristics of each separate stimulus. It is, therefore, difficult to discover in individuals a distinct type of imagery.<sup>1</sup>

5. . An individual's estimate of his own type of imagery does not necessarily describe the kind of imagery which can be voluntarily summoned most vividly, nor does it anticipate the kind of imagery which will predominate in his solving of simple problems.

---

<sup>1</sup>Similar results have been noted by Colvin, S. S. *Op. cit.* p. 235, and by Fernald, M. R. *Op. cit.* p. 130.

## CHAPTER III.

### SITTING I.

The main purpose of this sitting was to familiarize the subject with the general procedure of the experiment. The following instructions were given to the subject: "This is an experiment in thought. I am going to ask you to solve a problem which I hope will be solved in some way which will be new to you. I will, therefore, ask you what you would do if you were in a certain situation which I think will be a novel one for you, or else I will ask for your opinion about some matter which I think you have not considered before. The problem will be simply stated so that you will not need to ask questions about it. As soon as you hear the problem, will you think it over until you reach a conclusion which seems new to you. Then you may signal to me as soon as you have reached this point."

The experimenter pressed her key when she started to give the problem, released it when she finished, described the subject's movements by her key, making use of the code,<sup>1</sup> and pressed the key again when she received the signal from the subject that he had reached a new conclusion. This record was made on the kymograph with a simultaneous record of the time in fifths of a second. The answers of the subject to the required questions<sup>2</sup> were recorded verbatim after each problem was solved. Five problems were solved by each of the ten subjects.

A study of the results of this sitting shows that the gross bodily behavior of each individual represented a characteristic mode of adaptation which he adopted toward all his problems. Subject D, for instance, after having heard the problem, invariable raised her right arm to the arm of her chair and rested her chin on it. Subject E, on the contrary, looked downward, raised her right hand to her chin, rubbing her chin almost continuously throughout her thinking. Subject C, however, looked downward before she said she was ready for the problem and remained comparatively motionless throughout. In a similar way, Subject F leaned his head

---

<sup>1</sup>See Appendix, p. 89.

<sup>2</sup>See above p. 17.



on his right hand supported by the arm of his chair and fixed his gaze on some spot before he signalled that he was ready for the problem. He then maintained this position, with slight change, throughout his thinking. In one instance, however, when he was not in this position at the time the problem was given, he adopted this attitude as soon as he had heard it. This gross bodily behavior was so invariably adopted by each individual that it seemed to have no bearing on the solving of the various problems, but was simply a bodily adjustment for the better concentration of attention.

Aside from these movements of adaptation, more minute movements which appeared more irregularly were observed by the experimenter, such as eye and lip movements, smiles, frowns, and nods. These movements seemed to be more relevant to the solving of each specific problem, as there was a certain degree of correspondence between them and the thought and imagery which were described in the subject's retrospections.

Another kind of movement was noted by the experimenter which seemed to have no noticeable connection with the thought going on at the time and which apparently played no part in adapting the individual for better concentration of attention. These movements were those of the larger muscles, such as: random movements of hands and feet, scratching the head, stroking the hair, etc. These have been termed "doubtful movements," as their connection with the thinking was not apparent.

A distribution of the different kinds of movement is shown in Table III where the reactions of all the subjects to all of their problems are divided into five-second periods, and the average number of each kind of movement appearing in each period is given.

The results given in Table III show that in general the number of movements tended to diminish as the thought became prolonged. The more minute movements which seemed to be relevant to the thinking were those most frequently resorted to. Doubtful movements were less frequent in their appearance than movements of adaptation and they failed to persist throughout the thinking as did the relevant movements and those of adaptation.

In order to bring to light more clearly the relation of those different kinds of movement to the development of thought,

TABLE III

5 sec. Periods	No. of Reactions	Ave. No. of M'v'ts.	Ave. No. of Adapt. M'v'ts.	Ave. No. of Doubt. M'v'ts.	Ave. No. of Rel. M'v'ts
1	50	1.9	.58	.40	.92
2	49	1.6	.34	.43	.79
3	41	1.5	.39	.29	.90
4	36	1.1	.33	.27	.55
5	32	1.09	.28	.20	.56
6	27	1	.47	.07	.51
7	21	1.1	.47	.33	.33
8	20	1.1	.45	.25	.50
9	19	.7	.21	.26	.31
10	18	.9	.33	.16	.44
11	14	1.07	.21	.14	.71
12	13	1.1	.38	.23	.53
13	11	.5	.09	.09	.36
14	9	.8	.33	.11	.44
15	6	1.	.16	.33	.50
16	5	.4	.40		
17	4	.2			.25
18	4	.5	.25	.25	
19	4	.7	.25		.50
20	4	.5	.50		
21	4	1.	.75		.25
22	4	1.7	.75		1.
23	1				
24	1	1.	1.		

Table IV is given below in which the reaction time of each problem is divided into four parts<sup>1</sup> and the extent of the different kinds of movement in each quarter is shown.

TABLE IV

Quarters of Reaction	M'v'ts. of Adaptation	Doubtful Movements	Relevant Movements	Total Number of Movements
I	28.6%	26.2%	45.2%	164
II	34.8	19.6	45.6	132
III	26.9	17.3	55.8	104
IV	29.7	18.	52.3	94

From these results it is apparent that movement decreased as thinking progressed. Relevant movements were throughout most conspicuous, especially in the second half of the thinking where they were more numerous than both of the other kinds combined.

A closer scrutiny of the kymographic records showed that, in spite of the greater number of movements found in the first quarter, forty-three of the fifty reactions showed a perceptible pause immediately after the presentation of the problem where no movements were observed by the exper-

<sup>1</sup>This reaction time does not include that of the presentation of the problem.

imenter.<sup>1</sup> The average length of this pause was 2.7 seconds, and since the average reaction time for solving the problems was 34.4 seconds, making the average length of the quarters only 8.5 seconds, this initial pause was an obvious one.<sup>2</sup>

During this pause it is possible that no movements at all were made by the subject, or that there were movements which could not be observed by the experimenter. In thirteen of the forty-three reactions in which this pause appeared, visual images were reported to have followed the preception of the problem; in eight, verbal images; in three, both visual and verbal; and in four, visual with other kinds of concrete imagery. In ten of these reactions the subjects reported that they "thought about" various things but they failed to describe their states of mind any further. In five reactions, however, the mental states immediately following the presentation of the problem were definitely described as imageless: the recollection of a preceding problem as a "total not split up," a "feeling of amusement," a "feeling of uncertainty," a "search for the meaning," and an "Einstellung towards the problem."

In the seven reactions in which this initial pause was not apparent, visual imagery was reported to have followed the perception of the problem once, while no accompanying eye-movement was observed by the experimenter. Both verbal and visual imagery appeared once, when accompanying lip movements were detected by the experimenter. In two reactions the subjects did not describe their first states of mind, and in the remaining three reactions imagery was reported absent, in its stead being a "feeling of novelty," of "being at a loss," and inner speech. From this conflicting evidence it is, therefore, impossible to conclude whether or not, during this initial pause, movements, too slight to be observed, accompanied the mental states which were experienced.

Another pause which was found to be characteristic of a large number of reactions, forty-one, immediately preceded the final signal given by the subject. The average length of this pause was 3.2 seconds, slightly longer than the initial pause. This pause is significant as the subject was told to signal to the experimenter as soon as some novel solution

---

<sup>1</sup>No pause less than two-fifths of a second was included.

<sup>2</sup>A similar pause was present in forty-one of forty-nine reactions in which the subjects failed to reach novel conclusions.



came to mind. The truly productive phase of thinking, therefore, very likely preceded this final signal and was apparently usually characterized by a cessation of noticeable movements.<sup>1</sup>

In describing the novel part of their thinking in these forty-one reactions, the subjects reported the presence of visual images in ten instances, verbal images in seven, both visual and verbal in twelve, a combination of concrete imagery in two, and auditory imagery in one. The novel part of four reactions the subjects were unable to describe. Five solutions, however, were definitely described as imageless: two were "flashes," the others "a consciousness of forward movement," "non-sensory but clear," "no terms at all, merely a feeling of plausibility."<sup>2</sup>

In the nine reactions in which this final pause was not apparent two novel solutions were described as verbal, the experimenter detecting accompanying lip movements in one instance. Two were described as containing both verbal and visual imagery, one of which showed accompanying eye movements. Three novel solutions the subjects were unable to describe, one of which showed the presence of lip movements. Finally, two were characterized as "ideas," neither of which were accompanied by "relevant" movements which could be detected. Again evidence is too conflicting to be able to state the likelihood of the subject to make incipient movements during this final pause when he apparently "stopped to think."

Imagery was reported to have been present in every one of the fifty reactions. In twenty-one reactions it was said to have been present throughout the thinking, while in the remaining twenty-nine reactions it came and went. Visual imagery alone was reported present in ten solutions, while verbal imagery appeared alone in but two solutions. The latter was usually, in thirty-six of the fifty reactions, accompanied by concrete imagery. Visual imagery, therefore, seemed to have been more important to the solving of these problems than did verbal imagery.

Many states of consciousness, other than those previously described, were definitely reported to have possessed no

<sup>1</sup>This pause was not as conspicuous in reactions in which the subjects failed to reach novel conclusions. Twenty-one of forty-nine reactions failed to show this pause.

<sup>2</sup>Cf. Woodworth, R. S., *Imageless Thought*. J. of Phil. Psych. etc. 1906 III. p. 706.

imagery. In four instances, periods of "blankness" were described, also alluded to as "gaps in the progress" and "periods of waiting for something to happen." The remembrance of an event was once described as "knowing," independent of imagery. One subject reported that though she could usually summon images of concrete things, in the thinking out of her problem she merely "thought of cream, cereal, and lots of staple products." Another subject said that an idea, though distinct, was not necessarily related to imagery, as it came before the images. The acceptance of an alternative was described by another as a kind of feeling, rather than sensory. A "feeling of direction" was also differentiated from imagery. Finally, some thinking was described as abstract: "the ruling out of soldiers was abstract, not in terms of imagery." In no case was the term "Bewusstseinsalage" used by any subject and in but one instance was the term "imageless thought" resorted to.<sup>1</sup> The importance of these imageless states to thinking is difficult to determine from the above evidence.

The following table, Table V, shows the amount of the different kinds of movement characterizing the instances of thinking, grouped according to their length of reaction time.

TABLE V

Reaction time	No. of Reactions	Av. No. M'v'ts.	Av. No. Adapt. M'v'ts.	Av. No. Doubt. M'v'ts.	Av. No. Rel. M'v'ts.
-10 sec.	9	2.7	.55	.22	2.1
10-20	9	3.5	1.1	1.	3.5
20-30	11	7.5	1.4	1.8	4.2
30-40	2	10.5	2.	.5	8.
40-50	5	8.2	3.4	3.4	1.4
50-60	3	11.3	5.3	1.6	4.3
60-70	5	20.2	5.4	3.6	11.2
70-80	2	13.5	6.	6.	1.5
80-90					
90-100					
100-110	3	27.	9.6	7.	11.6
110-120	1	29.	12.	1.	11.

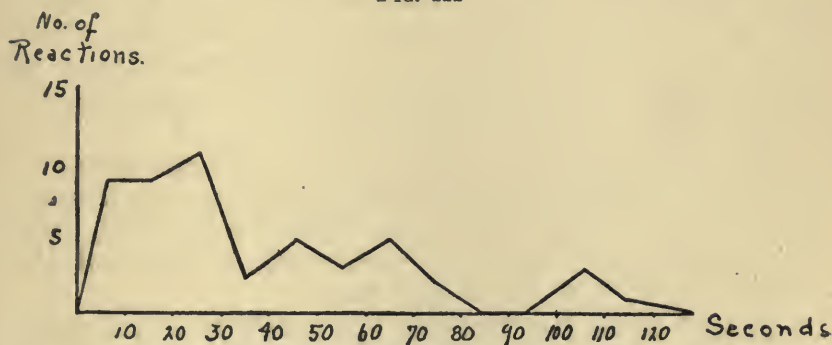
These results show that the number of movements of adaptation which were resorted to tended to vary in proportion to the length of reaction time. Relevant and doubtful movements, on the other hand, failed to show such a steady increase. The frequency of the doubtful movements resembled that of the movements of adaptation more than that of the relevant movements. The latter, however, tended to be more

<sup>1</sup>Subject I; his consciousness of "tablets" he reported to have been "like imageless thought." (Problem 12)

numerous throughout than both of the other kinds of movement.

When the distribution of the reactions is presented in graphic form, giving the frequency of reactions for each division in length of time required for thinking, Fig. III, the trimodal appearance of the curve makes it possible to describe some of the reactions as "rapid" (less than thirty seconds in length), "moderate" (between thirty and eighty seconds long), and "slow" (over 100 seconds long). The majority of reactions, twenty-nine, are shown to have been rapid.

FIG. III



FIFTY THOUGHT REACTIONS  
(Arranged according to time required for thinking.)

A comparison of these three classes of reactions, in respect to the average number of the different kinds of movement found in each, is shown in the following table, Table VI.

TABLE VI

Kinds of Reac.	No. of Reac.	Av. Time of Reac.	Av. No. of M'v'ts.	Av. No. of Adapt. M'v'ts.	Av. No. of Doubt. M'v'ts.	Av. No. of Relevant M'v'ts.
Rapid	29	16.1 sec.	5.5	1.06	1.06	3.3
Moderate	17	56.1	13.1	4.4	3.1	5.5
Slow	4	111.3	27.5	10.5	5.	12.

From these results it is apparent that the thinking, whether rapid or slow, was characterized more by minute relevant movements than by movements of adaptation or of doubtful character; movements of adaptation, however, varied more concomitantly with the length of the thought process. In moderate and slow thinking, therefore, the subjects repeatedly adapted themselves for better concentration of attention and frequently resorted to movements of the



limbs, the relations of which to the thinking were doubtful.

As regards imagery, twelve of the twenty-one reactions in which imagery was reported present throughout the thinking were instances of rapid thinking. Only one case of slow thinking showed this characteristic. Moreover, eight of the ten reactions in which visual imagery alone was present were quick reactions. No slow reactions contained merely visual imagery, for in every instance verbal imagery was also present. On the other hand, five of the seven reactions, in which the actually novel part of the thinking was described as definitely imageless, were instances of rapid thinking, and all of the reactions in which the subjects were unable to describe this part of their thinking were also quick reactions. Consequently, one may conclude that rapid thinking may be characterized by a predominance of concrete imagery or by states of mind lacking in imagery of any sort.

This classification of reactions into the rapid, the moderate, and the slow brought to light individual differences in the subjects. All of the reactions of A, C and H were rapid, none of the reactions of D and E were rapid, and none of the reactions of B, F, G, and J were slow.

In spite of the fact that this was merely an introductory sitting, its results, which are not contradicted by those of later sittings, seem to justify the following conclusions:

I. In thinking out problems individuals take certain characteristic bodily attitudes of adaptation for the better concentration of attention.

II. Movements are less apparent as thought develops.

III. In thinking, the most frequent kind of movement is of a minute sort: movements of eyes and lips, scowls and nods, which seem to be relevant to the changing content of consciousness.

IV. Immediately following the presentation of the problem, and also just preceding the termination of the thinking when some novel conclusion has presented itself, the thinker usually becomes noticeably quiet.

V. Either concrete visual imagery or verbal imagery may be the only kind of imagery present in the solving of a problem or in the actual production of a new conclusion. Neither kind of imagery alone, however, characterizes prolonged thinking.

VI. Imagery in thinking usually comes and goes, al-

though it is frequently present throughout the thinking, especially in quick reactions.

VII. Conscious states may be experienced which cannot be described in terms of concrete or verbal imagery. These may characterize the actual production of a new conclusion, especially in instances of rapid thinking.

VIII. Rapid thinking is characterized by either an abundance of concrete imagery, or by the prevalence of conscious states imageless in character.

## CHAPTER IV.

### SITTING II.

In this sitting the following instructions were given to the subject: "I am going to ask you to solve problems in the same way that you did before, with the exception that you are to press the key on the arm of your chair whenever you are conscious of the presence of images during your thinking, releasing the key when the imagery disappears. I will ask you to do this while I am giving you the problem if you happen to have imagery during that period."

It was the purpose of the experimenter to secure primarily a quantitative measure of the amount of imagery found in thinking. The kymograph was, therefore, arranged to record simultaneously the time in fifths of a second, the experimenter's observations of movements as in Sitting I, and the reactions of the subject's key showing the presence and absence of imagery.<sup>1</sup> A verbatim record of each subject's answers to the required questions for each of the five problems was taken as in Sitting I.

There is no doubt but that pressing the key to indicate the presence of imagery offered some distraction to the subject. Seven problems of Sitting I were given in Sitting II to different subjects. Six of these had a longer average reaction in the second sitting than in the first. But since seven of the ten subjects kept the same general record for speed which they had in Sitting I, and since only one reported the presence of distraction after the first two or three trials, it seems apparent that the subjects adapted themselves to the situation very readily.<sup>2</sup> In fact, it seemed to have become a habit for two who, in Sitting IV, unconsciously reached for the key when they were conscious of imagery.

The following table, Table VII, shows the average number of each kind of movement and the average time of imagery in seconds found in every five-second period of all the reactions of this sitting.

<sup>1</sup>For typical kymographic records, See Fig. VIII, Appendix p. 96.

<sup>2</sup>Subjects A, C, and D were given extra sittings in which they practiced solving problems with this use of their key. They all reported that although the act never became automatic, it could be used without interference.



TABLE VII

5 sec. Periods	No. of Reactions	M'v'ts.	Av. No. of Adapt. M'v'ts.	Doubt. M'v'ts.	Rel. M'v'ts.	Av. Time of Imagery sec.
1	50	1.8	.58	.20	1.02	2.20
2	49	1.1	.55	.31	.40	2.89
3	45	.9	.39	.13	.44	2.96
4	43	1.2	.55	.15	.67	2.89
5	40	.9	.62	.22	.62	2.75
6	35	1.1	.35	.26	.52	2.82
7	34	.9	.15	.34	.53	3.50
8	32	.9	.34	.28	.46	3.12
9	28	.9	.47	.17	.57	3.03
10	23	.9	.13	.26	.59	2.68
11	22	.7	.09	.42	.27	2.76
12	21	1.3	.61	.22	.72	3.68
13	18	.8	.37	.06	.54	3.28
14	16	1.	.24	.18	.54	2.90
15	11	1.4	.36	.27	.81	2.94
16	11	.8	.18	.27	.36	2.69
17	9	1.1	.11	.55	.44	2.66
18	9	.8	.11	.22	.55	2.26
19	9	1.1	.22	.66	.55	2.08
20	9	.7	.22	.22	.33	2.48
21	9	.7	.11	.55	.11	2.35
22	7	.8	.28	.28	.28	1.88
23	5	1.2	.40	.30	.60	2.88
24	5	1.4	.40		1.	1.00
25	5	.2	.20			1.20
26	3	1.3			1.33	1.93
27	2	1.			1.	2.50
28	2	3.	1.	1.	1.	2.50
29	1	2.	1.		1.	.60
30	1					
31	1					
32	1				1.	
33	1	1.			1.	2.
34	1	2.	1.		1.	
35	1	2.	1.			
36	1					
37	1					
38	1	1.		1.		
39	1					
40	1	1.			1.	
41	1	1.		1.		

The results of this table show characteristics similar to those given in Table III of Sitting I. In general, the number of movements tended to decrease as the thought became prolonged. Relevant movements were those most frequently resorted to. Again, the doubtful movements were least frequent, failing to appear in the latter part of long periods of thinking, where movements of adaptation and relevant movements continued to appear.

The above results also show that imagery as well was not always present in the latter part of the longer periods of

thinking. It is noteworthy, moreover, that the amount of imagery was not greatest in the first part of the thinking, in spite of the fact that the kymographic records frequently showed the presence of imagery during the presentation of the problem. On the contrary, the amount of imagery tended to increase after the first few moments of thought.

Table VIII, which follows, gives the extent of each kind of movement in terms of the percentage of the total number of movements recorded for each quarter of all the reactions, as in Sitting I, as well as the percentage of the total time of that quarter in which imagery was present.

TABLE VIII

Quarters of Reactions	Movements of Adaptation	Doubtful Movements	Relevant Movements	Total No. Movements	Amt. of Imagery
I	39%	14.2%	46.8%	182	46.8%
II	34.6	14.6	50.8	150	53.2
III	21.1	27.8	51.1	151	56.
IV	25.	30.	45.	140	55.6
				Total 623	Av. 52.9

These results agree closely with those of Table IV, Sitting I, and with those of the preceding table. Movements in general tended to decrease as the thinking progressed. Moreover, relevant movements were those most frequently present as before and they remained the most prominent kind of movement in each quarter of the thinking. This sitting, however, showed an increase in the number of doubtful movements, especially in the second half of the thinking, as compared with those recorded in Sitting I.

As in Table VII, imagery is found to have increased in amount after the first part of thinking. Its development was more closely paralleled by the changes in relevant movements than in any other kind of movement.

Again, as in Sitting I, in spite of the great number of movements in the first quarter, a larger majority of the reactions, forty-five, showed the presence of a pause following the presentation of the problem. The average length of this pause was 3.5 seconds, slightly longer than the initial pause of Sitting I. The average time for solving the problems of this sitting, however, was longer than that of the previous sitting, 55.8 seconds, probably due to the pressing of the key,<sup>1</sup> so

<sup>1</sup>See above p. 36.

that the length of this pause might be considered proportional to that of Sitting I.

In twenty-four of the forty-five reactions in which this pause appeared, the kymographic records showed that imagery, either concrete or verbal, accompanied the pause. In five of the remaining twenty-one reactions, where no imagery was recorded during this initial pause, states of consciousness which were definitely described as imageless followed the presentation of the problem. These were: a "feeling of recognition," a "feeling of being directed toward the ocean," a "feeling of being at a loss," the "thought that this question was related to the preceding one," and "the grouping of the problem as a whole," which was later followed by imagery.

In two of the five reactions where no initial pause was present the kymographic records showed that imagery followed the presentation of the problem. In one instance where visual and verbal images were described both eye and lip movements were detected by the experimenter. In the other instance, visual imagery was present, but it was described as fleeting, and no eye movements were noticeable. It is again difficult to decide, as in Sitting I, whether this pause is due to a lack of movement or to the presence of movements too fine to be detected by an onlooker.

The pause which preceded the subject's final signal was found to have characterized a large proportion of the reactions, thirty-nine, in this sitting as in Sitting I. The average length of this final pause was 4.5 seconds. Mental imagery, either concrete or verbal, was present during this pause in twenty-seven of these reactions, according to the kymographic records. Of the twelve remaining reactions, some of the conscious states which were reported as imageless were described in the following ways: "abstract thinking, experiencing relations which couldn't be analyzed," "proof" which followed the awareness of a possible solution, "awareness of similarity, of limitations, and of differences," and the "thought that the solution was good enough as long as no other idea was coming up."

Of the eleven reactions which showed no final pause, seven had imagery preceding the giving of the final signal, four of these had verbal imagery, in all of which cases lip movements were detected. Visual imagery was reported in three



instances, but no accompanying eye movements were detected by the experimenter. The remaining reactions which were definitely imageless in this last part of the thinking, as shown by the kymographic records, were described as "release in tension," "the feeling that these points will do," and an analysis "to see if this is the best solution." Again, with evidence so conflicting, it is difficult to explain the presence of this final pause.

Some imagery was described in the retrospections of all of the fifty reactions with the exception of three,<sup>1</sup> those of Subject A who was herself surprised at the fact.<sup>2</sup>

In sixteen of the remaining forty-seven reactions the kymographic records showed the presence of imagery throughout the complete thought process. Two of these reactions contained merely visual images. Verbal imagery alone was present throughout but one reaction, being usually, in eleven reactions, accompanied by concrete images.

In nine of the forty-seven reactions which contained imagery the subjects reported that merely concrete imagery was present. It is noteworthy that eight of these nine reactions were instances of rapid thinking. One reaction alone was said to have contained only verbal imagery; this being usually, in thirty-seven reactions, accompanied by concrete images. In the actually novel part of thinking, on the other

---

<sup>1</sup>Cf. Lummis, B. A. *A Study of Thought Processes*, p. 76.

<sup>2</sup>Problem I. 8. What do you think is the best way to teach geography to children? Ans.: I was at first at a loss, for I have no interest in geography. I knew I had talked about teaching geography in an objective way before. I had a feeling of knowing I remembered it. I turned to my personal experiences after I found the objective method was not new. Then I thought of telling them about my travels. I accepted this as something new. No imagery. Time 10.4 seconds.

Problem I. 9. During what years is it the best time to teach geography to children? Ans.: I thought first that this problem was connected with the preceding one. I did not know how to answer it. Then I thought it was better to begin with little children. It came so quickly, I don't know why I said this, perhaps because I had decided before to tell them about my traveling. No imagery. Time 6.6 seconds.

Problem II. 8. What would be the probable consequences if the sea could be owned by individuals as the land is? Ans.: I had a tendency of being directed toward the ocean between Europe and America. I wondered how any one would use it. Then I thought of fishing, but this wouldn't have serious consequences. It would be the same between individuals as between governments now. The problem would arise in regard to traffic, getting from one place to another. This problem would be solved in the same way that it has been for the railroads. I was satisfied that this relationship was new. Thinking was abstract, nothing specific. I had an experience of relations that couldn't be analyzed. No imagery. Time 36 seconds.

hand, verbal imagery predominated. It was present alone in eighteen solutions, while concrete imagery alone appeared fourteen times. Eleven solutions presented themselves in terms of both concrete and verbal imagery, while seven solutions could not be described in terms of either.

A study of the correspondence of movement and imagery as shown by the kymographic records brought to light the following facts: the majority of periods in which imagery was present (57.7 per cent) were accompanied by movements; 41.3 per cent of this movement was relevant, 38 per cent doubtful, while the remaining 20.7 per cent of the movements were for purposes of adaptation. When the periods of verbal imagery could be differentiated from those of concrete imagery by comparing the kymographic records with the subject's retrospections, 47 per cent of these periods were accompanied by noticeable lip movements. When the periods of visual imagery were singled out, 37.2 per cent were found to have been accompanied by perceptible eye movements. When of these the memory images could be differentiated from the imagination images,<sup>1</sup> more memory images were accompanied by eye movements (48.1 per cent) than was the case with the imagination images (14.3 per cent).<sup>2</sup>

Throughout the fifty reactions of this sitting there were 147 periods which the kymographic records showed to be imageless, breaks in the flow of imagery. During many of these (42.1 per cent) the subjects reported a feeling of progress. These were times when they "thought," they "stopped to analyze," or they "went over the entire field." Other imageless periods (34.6 per cent) were described as not accompanied by such a feeling of progress, being passive rather than active. During these periods the subjects "felt blank," were "receptive," "waiting for something to come." The characteristics of the remaining 23.1 per cent of these imageless states were not described by the subjects.

A study of the passive imageless states showed that the majority (64.7 per cent) were not accompanied by movement

<sup>1</sup>Images which were novel to the subjects.

<sup>2</sup>Perky, C. An Experimental Study of Imagination, *Am. J. of Psych.* 1913, XXI. p. 451. Martin, L. J. Die Projektions Methode u. d. Lokation visuelles u. a. Vorstellungsbilder. *Zsch. f. Psych.* 1912, LXI. p. 407, and Clark, H., Visual Imagery and Attention. *Am. J. of Psych.* 1916, XXVII, p. 490, have noted the accompaniment of eye movements with visual memory images. The evidence about the relation between ocular movements and visual imagination images is conflicting.



of any sort. Of the movements which did occur, 43.1 per cent were movements of the hands and feet, 31.3 per cent were lip and eye movements, and 25.4 per cent were those of bodily adaptation.

The active imageless states, on the other hand, were usually accompanied by movement (72.6 per cent). Thirty-five per cent of these movements were those of the lips and eyes, while 10 per cent were hand and foot movements.

These imageless periods constituted practically half of the time given to thinking, the subjects being conscious of imagery only 52.9 per cent of the time. A detailed description of these states would be exceedingly difficult, not only because of their variability, but also because the subjects frequently were unable to describe them, even when they were keenly conscious of their existence.

The following table, Table IX, shows further characteristics of thinking similar to those noted in Table V, Sitting I.

TABLE IX

Reac. Time	No. of Reactions	M'v'ts.	Av. No. of Adapt. M'v'ts.	Doubt. M'v'ts.	Rel. M'v'ts.	Av. Time of Imagery sec.
—10"	5	2.4	.6	.8	1.	2.2
10—20	5	5.8	.8	.6	4.4	12.8
20—30	6	4.3	1.6	.2	3.4	14.9
30—40	6	7.8	2.1	1.1	4.5	25.2
40—50	6	12.1	5.3	.6	6.3	21.4
50—60	4	13.5	4.7	6.	2.7	33.6
60—70	2	6.5	2.5	2.5	1.5	16.1
70—80	5	14.6	4.6	2.6	7.4	43.3
80—90	2	16.	6.5	6.5	3.	73.9
90—100						
100—110	2	20.	8.	9.	3.	37.4
110—120	2	25.	10.	3.	12.	23.5
120—130	2	41.	8.5	7.5	25.	7.6
130—140	1	35.	6.	3.	26.	2.2
140—150	1	15.	5.		10.	140.
190—200	1	36.	11.	8.	17.	4.6

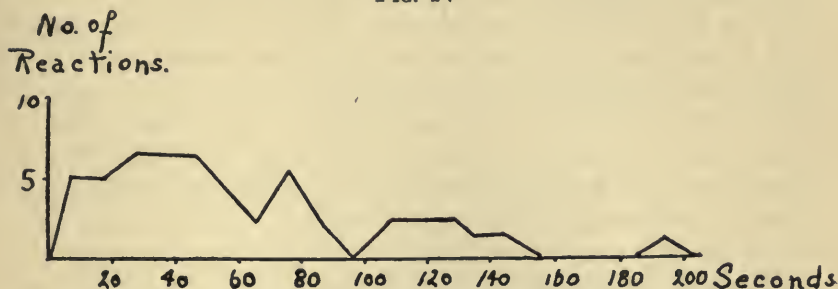
As before, the average number of each kind of movement for each division in length of reaction time is shown, while the last column gives the average time of thinking during which imagery was present. The number of movements of adaptation, as in Sitting I, was proportional to the length of reaction time, while the relevant and doubtful movements failed to show such a steady increase. Again the relevant movements were more numerous than both the doubtful movements and the movements of adaptation. The frequency of the doubtful movements resembled that of the movements of adaptation,



although there was no correspondence between the variabilities of each. As before, the amount of imagery did not particularly characterize the first few seconds in thinking, nor was it increasingly prominent in the longer reactions.

The following graph, Fig. IV, shows the distribution of all the reactions of this sitting according to the length of time required for thinking.

FIG. IV



## FIFTY THOUGHT REACTIONS

(Arranged according to time required for thinking)

As in Sitting I the distribution of these reactions makes it possible to classify them as "rapid," (less than 60 seconds,) "moderate" (between 60 and 100 seconds,) and "slow" (over 100 seconds in length). Here again, as in Sitting I, the majority of reactions, thirty-two, are found to have been rapid.

TABLE X

Kind of Reaction	No.	Av. time sec.	Av. no. M'v'ts.	Adapt. M'v'ts.	Av. no.		Av. time sec.
					Doubt. M'v'ts.	Relevant M'v'ts.	
Rapid	32	29.5	7.5	2.4	1.6	3.7	18.1
Moderate	9	72.7	13.1	4.5	3.4	5.1	44.
Slow	9	129.6	28.6	8.3	5.5	14.7	53.2

These results of Table X are similar to those of Table VI, Sitting I. Again the predominance of relevant movements, as compared with both doubtful movements and those of adaptation, is apparent. As already pointed out in Table IX, imagery is not conspicuous in either moderate or long reactions. There is, in this table, no noticeable similarity between the amount of imagery and the number of relevant movements as was apparent in Table VIII.

In spite of the fact that imagery is shown in the above tables not to have particularly characterized very rapid think-

ing, the majority of the kymographic records which showed the presence of imagery throughout the thinking, twelve of the sixteen, were quick reactions. Also, as in Sitting I, the majority of the reactions in which visual imagery was the only kind present, five of the six, were instances of rapid thinking. The slow reactions contained verbal imagery with the visual. On the other hand, the three reactions which were described as imageless belonged to this group of quick reactions; in fact, they were especially quick for the group, being 6, 10, and 36 seconds in length respectively. Again, as in Sitting I, of the majority of instances where the actual ly novel part of the thinking was definitely described as imageless, four of the seven were quick reactions. It is therefore true, as before, that rapid thinking was either characterized by a great deal of imagery or by conscious states which could not be described in terms of imagery.

Some individual differences in rapidity of thinking were found in this sitting, as in Sitting I. Subjects A and H were quick in all of their reactions, D and J were slow in none, and none of the reactions of I were quick. Subject C lost her former record of always thinking rapidly, D and E showed themselves capable of rapid thinking, while B, F, and G, resorted to slow thinking but once.

Finally, the exact record of the time during which imagery was present, which was taken in this sitting, makes it possible to correlate thinking with imagery, giving the following results: the correlations of the fifty reactions in respect to amount of imagery and length of reaction time gave  $r = .45$ . Taking the 32 reactions which have been called "rapid," this correlation between amount of imagery and length of reaction time gave  $r = .55$ . With the 9 "moderate" reactions, this correlation was .34, while from the 9 "slow" reactions a correlation of  $-.15$  was found.

These correlations indicate that the thinking in general was apt to have been accompanied by imagery: the faster the thinking, the more apparent the imagery, and, moreover, the slower the thinking the less apparent the imagery. As two-thirds of the 9 slow reactions showed a negative relation between imagery and movement and but one-third of the instances of moderate thinking showed this characteristic, the length of reaction time which would be likely to give a 0 correlation, and thus show a lack of relationship between the

rate of thinking and imagery, lies between 85 and 106 seconds. These facts support the statements already made above.

The results of this sitting, as have been shown, have offered further evidence to support the conclusions which were drawn from the results of Sitting I. They also offer evidence upon which the following conclusions may be based:

I. A problem may be comprehended, thought out, and solved in terms which cannot be described as any kind of imagery.

II. Imagery tends to increase in amount as the thinking develops until the very last steps in thinking are reached, when it frequently is absent.

III. Imagery is seldom present continuously throughout the solving of a problem. On the contrary, it usually comes and goes, actually taking up little more than half of the time given to thinking.

IV. Periods of imagery are often accompanied by perceptible movements, the kind of movement being frequently relevant to the imagery described.

V. Imageless periods are more often described as active, being accompanied by feelings of progress, than as passive. Such states are physically as well as mentally active, while the passive states are more frequently mentally and physically passive.

VI. Rapid thinking is correlated more highly with imagery than are instances of more moderate thinking; while, after a certain point, the longer the thinking the less the amount of imagery experienced.



## CHAPTER V.

### SITTING III.

The results of Sitting II made it possible to determine the kind of movement which accompanied the periods in which imagery was present or absent, but they failed to show the length of time actually taken up by these movements. As this evidence was given in the case of the imagery but not of the movements, a complete correspondence between the two could not be determined. Accordingly, in Sitting III the subjects were given the same instructions as in Sitting II. The experimenter, however, instead of recording the kind of movement she observed, pressed her key as soon as she detected movement on the part of the subject and released it when the movement ceased. The kymographic record, therefore, showed the presence and absence of imagery, as indicated by the subject's key, together with the presence and absence of movement, as indicated by the experimenter's key, accompanied by the time record in fifths of a second. The experimenter also kept a verbatim record of the answers to the required questions of each subject for all five reactions, as in Sittings I and II.

The following table, Table XI, gives the average amount of movement and imagery in terms of seconds found in each five-second period of all reactions of this sitting.

The results given are very similar to those of Sittings I and II. In general, the number of movements tended to decrease as the thinking was prolonged. The relative lack of imagery at the start followed by a slight increase in amount as the thinking progressed, disappearing in the latter part of long reactions is again noticeable. Imagery is found to have been more often present than perceptible movements, the total time for the former being 1,593.8 seconds as compared with 1,139.4 seconds during which movement was perceptible. 52.8 per cent of the time taken by movement was accompanied by the presence of imagery, while 37.9 per cent of the time during which imagery was present was accompanied by noticeable movements. Movement and imagery were apt to be combined in the earlier rather than in the later parts of the thinking. The amount of time during which movement and imagery

TABLE XI

5 sec. Periods	Number of Reactions	Movement	Av. Time of Imagery	Imagery & Movement Combined
		sec.	sec.	sec.
1	50	2.5	2.04	1.3
2	49	2.1	2.6	1.2
3	48	2.2	2.8	1.3
4	46	2.1	2.9	1.2
5	42	1.5	2.9	1.1
6	42	1.9	2.8	1.2
7	39	1.5	2.9	.8
8	36	1.6	2.6	1.1
9	28	1.6	3.02	1.4
10	24	1.1	3.6	.8
11	21	1.3	2.7	.7
12	21	1.4	2.4	1.09
13	19	1.02	3.	.6
14	18	1.02	3.03	.8
15	17	.7	2.6	.2
16	16	1.01	2.01	.5
17	13	.5	2.1	.5
18	13	1.2	2.2	.4
19	10	1.1	2.5	.4
20	10	1.02	2.4	.3
21	8	.5	1.9	.07
22	7	.8	2.2	.7
23	7	.9	1.7	.4
24	5	.6	1.9	.2
25	5	.2	1.08	
26	4		1.3	
27	4	.2	1.2	
28	3	2.3	1.6	
29	3	1.2	1.6	
30	3	2.1	.13	
31	2	2.6		
32	2	1.5		
33	2	3.8		
34	2	3.7		
35	2	1.3		
36	2	2.5		
37	2	.4		
38	2	2.4	.4	.2
39	2	2.4		
40	2	.7		
41	1	3.4		
42	1	3.		
43	1	.4		
44	1	5.		
45	1	4.6		
46	1	.6		
47	1			
48	1	1.8		
49	1			
50	1			
51	1	1.6		
52	1			
53	1			
54	1	3.6		
55	1			
56	1	2.4		

TABLE XI—*Continued*

5 sec. Periods	Number of Reactions	Movement	Imagery	Ave. Time of Imagery & Movement Combined
57	1	1.8		
58	1	2.4		
59	1	4.		
60	1	4.2		
61	1			
62	1	4.		
63	1	1.6		
64	1	3.4		
65	1	4.		
66	1	4.6		
67	1	4.8		
68	1	2.8		
69	1	.2		
70	1	2.8		
71	1	3.4		
72	1	1.2		
73	1	.8		
74	1	2.		
75	1	4.		
76	1	2.6		
77	1	2.		
78	1			
79	1			
80	1	3.2		
81	1	2.8		
82	1	3.2		
83	1	1.		
84	1			
85	1			
86	1			
87	1			
88	1	2.		
89	1	3.2		
90	1	1.2		
91	1	1.		
92	1			
93	1			
94	1			
95	1			
96	1	2.6		
97	1	2.6		
98	1		1.6	

were combined varied in proportion to the time of movement more than to the time of imagery.

Table XII which follows gives the distribution of the time given to imagery, to perceptible movements, and to the two combined throughout the different quarters of the reactions, in terms of the percentage of the total time of the reactions of this sitting.



TABLE XII

Quarters	Movement	Imagery	Move't & Imagery Combined
I	8.4%	11.4%	4.8%
II	7.9	12.9	4.1
III	8.8	12.6	5.5
IV	8.	13.5	4.4
	<hr/> Totals 33.1%	<hr/> 50.4%	<hr/> 18.8%

These results substantiate those of the preceding table as well as those of Sittings I and II in showing that there was a slight tendency for noticeable movements to decrease as the thinking terminated, and those of Sitting II in showing that imagery, on the other hand, tended to increase, although it was more conspicuous in the last quarter of the thinking in this sitting than in that of the previous sitting. The actual time given to noticeable movements was less than a third of the total time of thinking, while imagery was present a little more than a half of the time, as was the case in Sitting II. It is also apparent as before that the most marked increase in imagery appeared after the thinking of the first few seconds. The fact that the time during which movement and imagery were combined varied proportionally to the amount of movement more than to the amount of imagery present is again noticeable.

In this sitting the pause following the presentation of the problem was also frequent, being present in forty-one of the fifty reactions. The average length of this pause was shorter than in either Sittings I or II, being 1.23 seconds long, although the average reaction time for thinking was somewhat longer, 62.6 seconds. This may have been due to the fact that the subjects had become more familiar with the nature of the problems and so adapted themselves more readily to determine their solutions.

During this pause verbal imagery was reported to have been present sixteen times, visual images nine times, and conscious states described in neither of these terms five times. As the results of this sitting do not show the kinds of movement observed, it is impossible to learn what kinds of imagery were accompanied by the definite kinds of movement when this pause was not present.

The average length of the pause preceding the subjects' final signals in this sitting was 5.16 seconds, proportionally as long as the similar pauses in the two previous sittings. During

this pause, verbal images were reported to have been present nine times, visual images nine times, verbal and concrete imagery combined nine times, combined concrete imagery twice, and conscious states not described in terms of either verbal or concrete imagery six times. In the remaining fifteen reactions, where no pause was apparent and where movements occurred, these same kinds of imagery were present, but as the kinds of movement observed were not recorded by the experimenter, the exact correspondence between the movement and the imagery could not be discovered.

There were two problems in this sitting also which were solved without consciousness of imagery of any sort. They were reported by the same subject, A.<sup>1</sup>

In nine of the remaining forty-eight reactions, imagery was present throughout the thinking, seven of which reactions were of rapid thinking. An interrupted flow of imagery, therefore, characterized 78 per cent of the reactions. Eight reactions were said to have contained merely verbal imagery, five of which were quick reactions. Four reactions were described as having merely concrete imagery, all of which were examples of rapid thinking. The remaining 36 reactions were presented in both concrete and verbal terms. As for the actually original parts of the thinking, the subjects' retrospections showed that seventeen were present in merely verbal terms, ten in terms of concrete imagery alone, while fourteen included both the concrete and verbal elements. In six actually novel parts of the thinking, however, the subjects were not able to report the presence of either form of imagery: three appeared suddenly "like a flash" and so defied analysis, three were "facts" or "notions" impossible to describe, while three the subjects were doubtful about.

---

<sup>1</sup>Problem III. 8. If you should start now for the Battery, how could you possibly reach there in half an hour? Ans.: I had a feeling of being directed toward the subway, the map, and the real place: the harbor, the Statue of Liberty. I was not seeing, just being directed toward them. I thought of the elevated but rejected them almost before they came to mind. I thought: It takes longer than half an hour, I couldn't go. I was conscious of the fact that it was impossible, (a conclusion she had not thought of before). No imagery. T. 20 seconds.

Problem III. 9. What historic character would you choose to write a play about? Ans.: I first thought of Joan of Arc as she is my favorite character, but I thought Schiller had done it better. Then I thought of my general principle that I would never write anything that was not scientific, so I gave up any idea of writing a play. My mind jumped from one side to the other. I was conscious of the conclusion before I described it. No imagery. T. 10.2 seconds.



A study of the correspondence of movement and imagery as found in the kymographic records shows that when the actual time given to each was measured, imagery was less often accompanied by movement than not, 37.9 per cent. When the verbal images could be differentiated from the concrete images by comparing the kymographic records with the subjects' retrospections, it was found that the former were more often without noticeable accompanying movements than with, 56.7 per cent. When the visual images were similarly singled out, 40.2 per cent of them were found to have been accompanied by perceptible movements. When the memory visual images could be differentiated from the imaginary ones, they were found to have been more often accompanied by movement, 49.1 per cent, than were the imaginary images, 24.2 per cent. This correspondence between movement and imagery in this sitting, it must be remembered, simply refers to movement in general, as no record of the kinds of movement was kept by the experimenter.

The periods in which imagery was absent, 105 in all, were more often described as active than passive, 63.8 per cent being accompanied by a feeling of progress, while many of the others were described as "breaks," "cracks," or "blanks" in the thinking.<sup>1</sup> As in Sitting II, the majority of passive periods were not accompanied by perceptible movements, 68.4 per cent. In 82 per cent of the active states, on the other hand, some movements were noticeable.

The following table, Table XIII, gives variations in imagery and movement in proportion to the length of reaction time. These results, too, show characteristics similarly pointed out in Sitting II. Imagery again failed to characterize both the most rapid and the slowest thinking.

There was practically no correspondence between the presence of imagery and of perceptible movements, although both the imagery and the movement were frequently combined.

---

<sup>1</sup>Are these instances in which thought is "baffled"? Fox, C. *The Conditions which arouse Mental Imagery in Thought*. Rep. Brit. Ass. for Adv. Sci. 1913, LXXXIII. p. 688, and Betts, op. cit. p. 94, state that such conditions are especially favorable for arousing mental images.

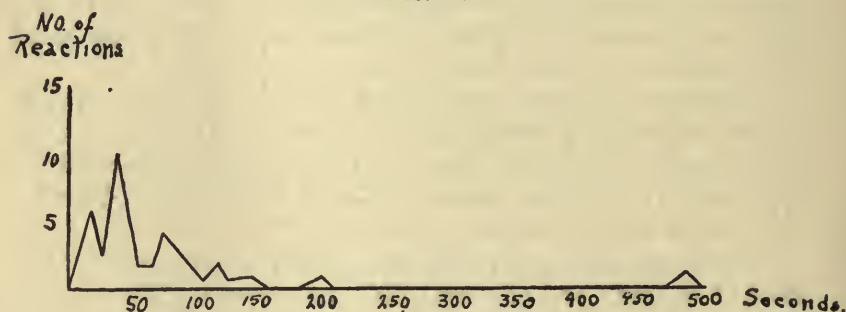


TABLE XIII

Reaction Time	No. of Reactions	Av. time of Movement	Av. time of Imagery	Av. time of Imagery and Movements Combined
-10 sec	2	3.1 sec.	1.3 sec.	.9 sec.
10-20	6	11.1	7.1	5.3
20-30	3	19.8	47.8	10.2
30-40	11	14.07	19.9	14.8
40-50	7	15.1	26.8	10.4
50-60	2	9.	29.4	3.2
60-70	2	25.1	25.8	15.5
70-80	4	42.2	35.	27.0
80-90	3	10.4	59.9	6.8
100-110	1	19.4	84.2	16.6
90-100	2	30.8	47.4	1.7
110-120	2	21.5	60.8	15.9
120-130	1	19.4	97.8	19.2
130-140	1	24.4	3.	1.2
140-150	1	23.	138.6	21.4
190-200	1	59.6	3.2	.8
490-500	1	64.8	5.	.4

The following graph, Fig. V, gives the distribution of the reactions of this sitting according to the length of their duration in terms of seconds.

FIG. V.



**FIFTY THOUGHT REACTIONS**  
(Arranged according to time required for thinking)

As in Sitzings I and II, it is possible to divide these reactions into three classes: "rapid" (less than 60 seconds), "moderate" (between 60 and 100 seconds), and "slow" (over 100 seconds). Again, as in the previous sittings, the majority of the reactions, 31, were rapid.

TABLE XIV

Kinds of Reaction	Number of Reactions	Av. time of Reaction	Av. time of Movement	Av. time of Imagery	Av. time of Imagery and Movement Combined
Rapid	31	33.05 sec.	15.4 sec.	17.3 sec.	9.6 sec.
Moderate	11	80.9	28.9	42.4	8.1
Slow	8	176.7	31.7	56.6	11.4

The results given in the above table show, as in Sitting II, the lack of a proportionate amount of imagery in instances of slow thinking. As for the instances of rapid thinking, it is true, as in Sittings I and II, that most of the reactions in which imagery was present throughout, 7 of the 9, were quick reactions. But it is also true, as in Sitting II, that all of the reactions which were described as wholly imageless were in this group. These imageless reactions were rapid even for this group, being 20 and 10.2 seconds in length respectively, but there were many reactions which were shorter and which contained imagery. All of the reactions whose actually novel parts were described as imageless were also instances of rapid thinking. On the other hand, all of the reactions which contained merely concrete imagery were of this group of quick reactions, and five of the eight which contained merely verbal imagery were also of this group. As in the previous sittings, therefore, rapid thinking may be characterized by a preponderance of imagery, especially of a single kind, or of imageless states. Finally, the combination of imagery and movement showed no variability proportional to the length of reaction time, to the amount of imagery, or to the amount of movement.

Individual differences in rapidity of thought are apparent in this sitting, as in the previous ones. A, G, and H were quick in all of their reactions, D, E, and J were not slow, while B was not quick. In comparison with the earlier results, A, H, and J still held their records of both previous sittings, while subjects D and E showed the same rate of thinking which they had in Sitting I.

Finally, from the results of this sitting, it was possible to secure certain correlations. By using the fifty reactions of this sitting, speed of thinking correlated with the amount of time given to imagery, to movement, and to imagery and movement combined as follows:

Speed of thinking with:	Imagery	$r = .21$
	Perceptible movements	$r = .19$
	Imagery and movement combined	$r = -.04$

When the quick, slow, and moderate reactions were in turn considered, their speed of thinking correlated with the amount of time given to imagery, to movement, and to both in combination, showing the following correlations:

Speed in rapid thinking with:	(31 reactions)
Imagery	$r = .59$
Perceptible movements	$r = .401$
Imagery and movement combined	$r = .32$
Speed in moderate thinking with:	(11 reactions)
Imagery	$r = .32$
Perceptible movements	$r = -.01$
Imagery and movement combined	$r = -.002$
Speed in slow thinking with:	(8 reactions)
Imagery	$r = -.41$
Perceptible movements	$r = .09$
Imagery and movement combined	$r = -.46$

With the fifty reactions of the sitting, the time of imagery which was recorded, and the time during which movements were perceptible showed a correlation of  $r = .209$ .

Finally, when the results of Sittings II and III were combined, it was possible to get the correlation between imagery and speed of thinking with the following results:

Imagery with:	
Speed of thinking (100 reactions)	$r = .17$
Speed of rapid thinking (63 reactions)	$r = .506$
Speed of moderate thinking (20 reactions)	$r = .35$
Speed of slow thinking (17 reactions)	$r = -.46$

The information given by these correlations is similar to that already submitted. In general, imagery and speed of thinking were not correlated to any degree, while the relationship between speed of thinking and perceptible movements was even less close. However, the more rapid the thinking, the more apt it was to have been accompanied by imagery. This is true in spite of the fact that two instances of rapid thinking were reactions which did not contain imagery. Although based upon but relatively few instances, the high minus correlation between slow thinking and imagery suggests the possibility that the slower the thinking, the less the imagery, a fact somewhat substantiated by the results of Sitting II. Imagery was not correlated to any extent with perceptible movements in general.

As four of the eleven instances of moderate thinking showed a negative relationship between the amount of imagery and the speed of thinking, while six of the eight instances of



rapid thinking showed this characteristic, it seems likely that those reactions between 95 and 105 seconds in length were those which showed no relationship whatsoever between imagery and speed of thinking. This reaction time resembled very closely that similarly obtained in Sitting II.

The correlations secured when the reactions of both sittings were combined were very much the same as those which had been given in each sitting, further corroborating the conclusions already stated.

The results of this sitting, as has already been shown, have offered further evidence to support the conclusions drawn in Sittings I and II. Moreover, they seem to present sufficient evidence upon which to base the following conclusion:

1. Perceptible movements accompany thinking during not more than one-third of its total time. They come and go at irregular intervals, with the exception of the pauses immediately after the presentation of the problem and just prior to the subject's final signal.<sup>1</sup>

---

<sup>1</sup>Already noted above, p. 34.

## CHAPTER VI.

### SITTING IV.

It was the purpose of this sitting to learn, if possible, how close the relationship between imagery and relevant movements could be. The following results of Sitting II, already noted, indicated that a possible connection existed between imagery and movement, such that one might expect that movements had been constantly accompanying the images but were in themselves too minute to be readily detected by an onlooker.

1. Of the 301 relevant movements recorded by the experimenter, 8.8 per cent were eye movements which accompanied visual imagery, while 10.8 per cent were lip movements accompanying verbal images.

2. 46.7 per cent of the periods where imagery was present were accompanied by relevant movements.

3. Of the 86 visual images, 32, or 37.2 per cent, were accompanied by noticeable eye movements, while of 104 verbal images, 49, or 46.1 per cent were found to have had accompanying lip movements.

No description of the "verbal images" which entered into the subjects' thinking has been attempted up to this point, as it was difficult for the experimenter, even with the help of the objective records, and, in many cases, for the subject to determine to what extent the imagery was in reality inner speech. The subjects were allowed to describe their verbal elements in any terms they wished,<sup>1</sup> and as six of them nearly all of the time, and all of them some of the time referred to it as "imagery,"<sup>2</sup> the experimenter has always referred to it as such up to this point, until an opportunity should arrive for differentiating between the two.

Various forms of distraction were introduced in this sitting in order to detect the degree of the connection between

---

<sup>1</sup>The following expressions were used: "vocal motor," "auditory motor," "auditory kinaesthetic," "auditory," "visual," and "verbal" imagery.

<sup>2</sup>Three subjects, C, F, and H, questioned as to whether or not some of their verbal images gave them kinaesthetic sensations. Subjects A, B, and C decided that some of their images were sensational, while Subject I admitted that he frequently talked to himself.

imagery and accompanying movements.<sup>1</sup> The visual distraction was a white bob swinging in a large arc over the head of the experimenter. The subject was asked to solve a problem which was likely to arouse visual imagery<sup>2</sup> while he gazed constantly at the swinging bob. The experimenter in the meantime recorded the failures of the subjects to follow this movement, pressing her key when the gaze wandered, and releasing it as soon as it was again fixed on the bob.

The auditory distraction was a bell rung from the moment the problem had been presented until the subject's final signal was given. The problem was selected as one which would be likely to require auditory imagery for its solution, in order to determine whether or not there would be interference between sensations and images of the same kind. During the solving of the problem the experimenter recorded the subject's movements, as in *Sittings I and II*.

The verbal distraction was the repetition aloud of a b c, a b c, a b c, etc., as fast as possible. The subject was given a problem which was likely to require verbal imagery for its solution. The experimenter recorded the presence of any hesitation or confusion on the part of the subject in pronouncing the letters.

Finally, each subject was given a problem which it was believed would require hand-motor imagery for its solution. The subject was asked to keep the thumb and forefinger of each hand in a simple rhythmical motion which he had previously practiced, so that it could be carried on automatically during the solving of the problem. The experimenter kept a record of any hesitation in this movement.

The kymographic records of this sitting showed the time in fifths of a second and the records of the experimenter, which varied with each problem as described above. The answers of each subject to the required questions were recorded verbatim, as in the previous sittings.

The results of the experiment with the visual distraction showed:

---

<sup>1</sup>A similar use of distractions has been made by Münsterberg, H., and Campbell, *Psych. Rev.* 1894, I. p. 441, Smith, T. L., *Am. J. of Psych.* VII. 1896, p. 459, Secor, W. B., *Am. J. of Psych.* 1897, XI. p. 232, Cohn, J., *Zsch. f. Psych.* 1897, XV. p. 161, Meumann, E., *Ueber Oekonomie und Technik des Lernens*, 1903, Segal, J. *Arch. f. d. ges. Psych.* 1908, XII. p. 124, Fernald, M. R., *Psych. Rev. Mon.* 1912, XIV. No. 58, and Pintner, R., *Psych. Rev.* 1913, XX. p. 129.

<sup>2</sup>See Appendix p. 93.



(a) Two subjects were completely distracted. They could not solve the problem.

F acknowledged that he was dependent upon visual fixation in his thinking. He was given another problem which he solved, but his eyes glanced aside from the swinging bob several times, during which he admitted his thought progressed.

A admitted that she could not attend to both the swinging bob and her imagery. She was given another problem and failed again. She then asked for a problem which would not require visual images. She was given one which could be solved by verbal imagery which she succeeded in solving without the aid of visual images. She was still distracted, however, and the experimenter's record showed that her gaze wandered several times. She was apt to use visual images in her thinking.

(b) Five subjects reported that they were distracted but still able to solve the problem. In all of these reactions visual images were said to have been present, and the experimenter's records showed that the eyes frequently failed to follow the bob. All of the subjects reported that they became aware that the motion of the bob interfered with their imagery so that they often wondered whether or not their eyes were following its motion. Some admitted that they had to voluntarily attend first to the one, then to the other, in order to accomplish the task. The eye movements recorded by the experimenter took place at periods in the thinking at which images were reported. Two subjects confessed that they were conscious of the side glances from the bob, but added that they were not done intentionally.

(c) Three subjects, H, I, and J, reported that they were not at all distracted. They said that they felt no disturbance, having images in their usual way. Subjects H and J were found to have used eye movements to the side at times when judging from the subjects' retrospections, visual imagery was reported to have been present. Of these eye movements the subjects seemed to have been wholly unconscious. The experimenter was unable to discover that Subject I looked away from the bob at all.

Fourteen visual imagination images were described in nine of the ten solutions. With nine of these images, the experi-

menter's records showed that the subjects' eyes had looked away from the swinging bob at points in the thinking which corresponded to the report of images in the subjects' introspections. With visual memory images, also, eight of the ten described were accompanied by such side glances.

The swinging movements of the bob would interfere with habits of visual fixation as well as with ocular movements relevant to the image itself. Two subjects definitely reported that their difficulty was due to an inability to fixate their gaze upon some one spot. This visual fixation did not seem to be related to any one special kind of imagery,<sup>1</sup> for with Subject F the images were those of memory while in the case of Subject C the images were novel and full of movement. On the contrary, this visual fixation characterized the general bodily adaptation of these subjects to the solving of all problems<sup>2</sup> and seemed to be merely a means of avoiding visual distractions.

The reasons for the side glances of the remaining seven subjects may then have been due to interference with slight movements accompanying the images.<sup>3</sup> But the difficulty definitely described by five subjects was that of control of attention. They stated that they were forced to attend alternately to the distraction and to their thinking. If the turning of the attention to the thinking, however, involved ocular movements other than those of following the swinging bob, such movements must have been inseparably connected with the contents of thought or with an unconscious habit of visual adaptation to the thinking.

The three remaining subjects who reported that they were not distracted must have adapted themselves automatically to following the swinging bob and so were able to attend more easily to the contents of their thoughts. But as two were found by the experimenter to have made glances to the side, this automatic motion must have interfered either with an unconscious habit of visual fixation or with incipient ocular movements accompanying visual images.

It cannot be said conclusively, therefore, that visual images were invariably accompanied by ocular movements. The

---

<sup>1</sup>Perky, C. W. *Op. cit.* p. 451, asserts that there is 100% visual fixation for "imagination" images, while Clark, H., *op. cit.* p. 491, states that there is lack of ocular movements when the usual conditions of perception are least reinstated by the image.

<sup>2</sup>See above, p. 28 ff.

<sup>3</sup>Cf. Clark, H., *op. cit.* p. 487.

above results seem to indicate that either visual fixation was necessary for better concentration of attention upon mental states, or slight eye movement accompanied the visual imagery which was present. In the case of Subject I, however, the possibility still remains that neither visual fixation nor ocular movements with his visual images were necessarily present.

The results of the experiment with the auditory distraction show that (a) four subjects were distracted, but still able to solve the problem. Subject E was emotionally disturbed, as she was wont to be by the sound of any bell. She solved the problem, however, and had, during her thinking, auditory verbal imagery. C and G were unable to call to mind certain names. C, however, succeeded when she tried another set of associations: the name coming in terms of auditory motor imagery.

(b) Two subjects reported the presence of concrete auditory imagery while hearing the sound of the bell.

(c) All subjects described their verbal elements as auditory, showing the presence of auditory imagery.

These results indicate that an auditory sensation did not necessarily inhibit the consciousness of imagery of the same kind. This is not surprising, as visual imagery was constantly being reported in the thinking which was carried on with open eyes.<sup>1</sup>

A study of the results of the experiment with the verbal distraction shows:

(a) Three subjects were completely distracted.

G reported that he "couldn't get anywhere" with the problem. He was given another which he succeeded in solving, as his verbal imagery "slipped in" during the periods when he stopped for breath.<sup>2</sup>

E reported that she felt "very self-conscious." She failed to solve two problems, but succeeded with a third, during which her verbal imagery appeared when she paused to breathe.

C was not able to reach a new conclusion for her first problem. She solved a second problem by letting her verbal images slip in between the repetition of the letters.

<sup>1</sup>Cf. Fernald, M. R. *Op. cit.* pp. 119 ff.

<sup>2</sup>Smith, T. H. *Op. cit.* p. 453 ff., Fernald, M. R., *op. cit.* p. 119, and Pintner, R., *op. cit.* p. 140 report similar results in the presence of verbal distractions.



(b) Two subjects, G and H, consciously abandoned their verbal images, solving the problem by means of concrete imagery.

(c) Three subjects were but slightly distracted. A and I confessed that their verbal imagery crept in when they paused during the repetition of the letters. F, on the other hand, so controlled his verbal images that they went simultaneously on "in his head."<sup>1</sup> The experimenter noted that he made a short pause in the repetition of the letters. Here the subject explained that he stopped "to give his points a hang."

(d) Two subjects were not at all distracted.

D used verbal imagery throughout, feeling no conflict between the two. She said she could "go on doing it forever." She paused three times for breath.

J did not report the presence of verbal images. The problem was solved in visual terms. His rate of repetition was noticeably retarded three times, however, although he said he was not distracted, as the repetition of the letters went on mechanically. His conclusion presented itself as a "notion."

These results seem to indicate that with some individuals articulatory movements are necessarily connected with their verbal elements in thinking. Their "verbal imagery" has, therefore, probably been, throughout the solving of their problems, inner speech. In the cases of D and F, however, the possibility still remains that they may have experienced actual verbal images.<sup>2</sup>

The experiment with the hand-motor distraction gave the following results:

(a) One subject was completely distracted. Subject A reported that if she had worked out the problem she "would have had to draw it." She was conscious of distraction continually. When a second problem was given she gained an "abstract knowledge" of it, but the solution did not come until she had stopped moving her fingers. A third problem she solved by talking aloud, finally stopping the movement of her fingers in order to complete the thought. She said that if she were given another kind of problem she could solve it in spite of the distraction.

(b) Three subjects were slightly distracted.

<sup>1</sup>Pintner, R., op. cit. p. 152 reports a similar change in the presence of verbal distraction.

<sup>2</sup>Reed, H. B. Inner Speech in Thought Processes. *J. of Exp. Psych.* 1, 1906, p. 372, admits this possibility.

B wanted to stop her finger movements in order to draw. She felt sensations of movement in her arms and hesitated several times in her finger movements. She solved the problem in verbal terms.

D also wanted to use her hand for drawing, but finally resorted to visualizing the movements.

G became confused and stopped moving his fingers. He continued in a moment, however, and solved the problem by concrete images.

(c) Six subjects reported that they experienced no distraction. C, F and I made use of hand-motor images, but reported that they felt no conflict between the movements made by their fingers and their motor images. All of their records, however, showed the presence of moments of hesitation in the movements of their fingers.

E, H, and J reported no hand-motor imagery in their solutions. In the case of E, however, there were three short pauses in her hand movements. H and J, on the contrary, showed no confusion or hesitation in their hand movements.

These results seem to indicate that with some individuals, motor imagery and incipient movements are inseparably connected. The possibility still remains in the cases of Subjects C, F, and I, however, that they may have experienced actual motor imagery apart from movement.

Summarizing the results of this sitting, the following conclusions seem justifiable:

I. Auditory and visual sensations do not inhibit the consciousness of imagery of the same kind.

II. With some individuals, visual images are inseparably connected with either visual fixation for better concentration of attention, or with slight relevant eye movements.

III. With some individuals, verbal imagery is inseparably connected with articulatory movements, showing that their "imagery" is in reality "inner speech."

IV. Hand motor imagery with some individuals is likewise inseparably connected with incipient hand movements.

V. There is no conclusive evidence that imagery is thus inseparably connected with movement for all individuals.

## CHAPTER VII.

### SITTING V.

On account of the lack of conclusive evidence that verbal imagery and articulatory movements are inseparably connected, this sitting was planned in which minute tongue and laryngeal movements could be measured by means of delicate apparatus.

Four subjects were examined in this way: A, who alone reported that she had solved problems without the presence of any kind of imagery, and who was found to have resorted to inner speech in Sitting IV; B, who had described her verbal elements throughout as images, but who showed confusion in the presence of verbal distraction; C, who was not always certain about the character of her verbal elements, but who was unable to solve her problem in the presence of a verbal distraction; and D, who carried on her verbal images with ease in the presence of the verbal distraction.

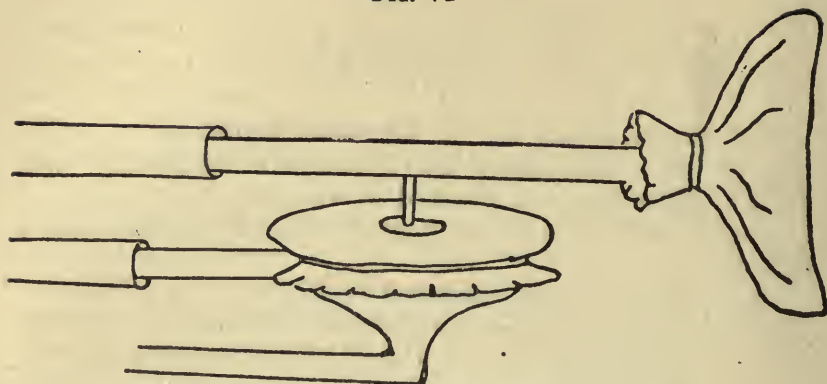
The work of this sitting was carried on in the Neurological Laboratory of the School of Physicians and Surgeons. The subjects were seated as comfortably as possible in a chair. They were given problems to be solved in the usual way, except that, during the thinking, records were made of their respiration, thoracic and abdominal; the changes in volume of the arm; their laryngeal movements, horizontal and vertical; and their tongue movements. The time record in fifths of a second was also kept, as well as a record from the experimenter's key which indicated the time of the presentation of the problem and the duration of the thinking. The kymographic records, therefore, showed eight separate tracings taken simultaneously.<sup>1</sup>

Sumner pneumographs were used to measure the changes in respiration. Lehmann's water plethysmograph measured the volume of the arm. To detect the laryngeal movements, two tambours were used, Fig. VI. The upper one was cup-shaped with a rubber membrane stretched over the top. This was fitted over the front of the throat and responded to the horizontal movements of the larynx. This tambour was fastened to the throat by means of a headpiece in such a way that

<sup>1</sup>See Figs. IX, X, X, XIX, Appendix p. 97 ff.



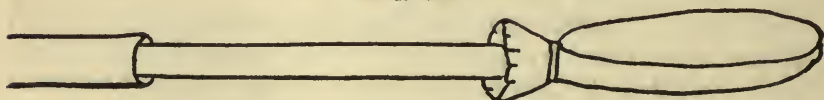
FIG. VI



Instrument for measuring laryngeal movements

the entire tambour moved up and down with the vertical movements of the larynx. In order to record these vertical movements, the stem of the cup-shaped tambour was attached at right angles to a second tambour which responded to these changes in position of the first. The tongue movements were recorded by means of a small flat frame, the size of the mouth, which was covered with a rubber membrane, Fig. VII., and placed in the mouth above the tongue.

FIG. VII



Instrument for measuring tongue movements

A continuous roll kymograph was used, run by electricity. The recording tambours were placed vertically on a frame above the kymographic paper. Their needles fell with the force of gravity upon the drum and recorded the movements of the tambours on the revolving sheet.

Under such conditions the subject was forced to sit perfectly still during the solving of the problems, thus eliminating all gross bodily movements. After the problems were solved the subjects' answers to the required questions were recorded verbatim, as in the previous sittings.

Subject A in her three problems reported the presence of verbal elements. She said they came "somehow" but she didn't know just how. The kymographic record showed the occasion-

al presence of inner speech in each reaction, as tongue movements were noticeable.<sup>1</sup> Laryngeal movements were occasionally recorded when the tongue was especially active in the first and last reactions. In the second reaction, however, the inner speech seemed to have been carried on by the tongue alone.

The volume of the arm, for subject A, showed a high pressure which was steadily maintained, not only during the period just before the problem was given, but during the presentation of the problem and the thinking, showing a slight gradual tendency to increase as the thinking progressed in the first and last reactions.<sup>2</sup>

The rapidity of the pulse, as recorded in the plethysmographic curve, showed no noticeable changes either before or after the presentation of the problem, or with the rise in the volume curve.<sup>3</sup> The amplitude of the pulse curve, however, showed a slight diminution with increase in volume.<sup>4</sup>

This subject's respiration records showed more marked differences between the periods of rest, of perception of the problem, of thinking, and of actual speech, than did any other records. Before the problem was presented her respiration was slow, deep, and regular. With the presentation of the problem it immediately became more rapid and superficial, although it still remained regular. This continued for a short period following the presentation of the problem<sup>5</sup> and then the breathing became deeper so that it resembled the normal

<sup>1</sup>Courten, H. C. Involuntary Movements of the Tongue. *Yale Psych. Studies* 1902, X. p. 94 reports the presence of tongue movements of which subject was not aware.

<sup>2</sup>Either constriction or dilation in the volume curve may accompany mental work according to Binet, A. and Courtier, J. *Les Causes d'Erreur dans les Expériences de plethysmograph.* *L'Ann. Psych.* 1895, II. P. 162 and Angell, J. R. and Thompson, H. B., *Organic Processes and Consciousness*, *Psych. Rev.* 1899, VI. p. 39.

<sup>3</sup>Regularity in the heart beat under sustained attention has been noted by Binet, A., and Courtier, J., *op. cit.* p. 158, and by Angell, J. R., and Thompson, H. B., *op. cit.* p. 48.

<sup>4</sup>This relationship between the volume curve and the amplitudes of the pulse record is affirmed by Binet, A., and Courtier, J., *op. cit.* p. 129, and by Lehmann, A., *Die körperlicher Aeusserungen psychischer Zustände*, but denied by MacDougall, R., *The Physical Characteristics of Attention*, *Psych. Rev.* 1898, III. p. 176.

<sup>5</sup>The rapidity and superficiality of the breathing in mental work has been reported by Binet, A., and Courtier, J., *op. cit.* p. 158, by Gamble, E. A. M., *Attention and Thoracic Breathing*, *Am. J. Psych.* 1905, XVI. p. 290, and by Stevens, H. C., *A Plethysmographic Study of Attention*, *Am. J. of Psych.* 1905, XVI. p. 468.

rate when the problem was solved.<sup>1</sup> When the subject started to speak at the end of the reaction, marked changes in the breathing curve were of course made.

These results can be contrasted with another reaction of this subject when she did not understand the question, tried to make something out of it, failed and finally gave up. In this reaction the tongue movements became less and less noticeable, the larynx showed no response and there was a marked fall in the volume curve, with a slight diminution in the height of the pulse curve.

The results given in Subject B's two reactions resemble greatly these of Subject A. Verbal elements were reported present in both reactions, the subject describing them as more auditory than motor. Her tongue movements were more minute than those of Subject A, but inner speech was at times noticeable in both reactions, the tongue movements in the solving of one problem being intense enough to have been accompanied occasionally by laryngeal movements.

In one reaction, when the subject thought of knitting, she reported the presence of a motor image of the attitude of knitting. At this point in the plethysmographic curve, a slight hand movement was perceptible, simultaneously with inner speech involving both tongue and larynx.

The volume curve in the records of this subject showed a slight gradual increase accompanied by a slight increase in rate of heart beat.<sup>2</sup> There was no perceptible decrease in the height of the pulse record as in the case of Subject A.

Subject B's respiration record was similar to those of Subject A.

The results of Subject C's five reactions also greatly resembled those of Subject A. Verbal elements were reported present in all reactions, and described as "more sensory than before" in the solving of the last two problems. This difference was noticeable in the kymographic records. As in the case of Subject B, Subject C's tongue movements were minute. Laryngeal movements were hardly noticeable except in the last two reactions. Inner speech was noticeable in the tongue

---

<sup>1</sup>Regularity of respiration accompanying periods of special concentration of attention has been reported by MacDougall, R., *op. cit.* p. 176, and by Angell, J. R., and Thompson, H. B., *op. cit.* p. 48.

<sup>2</sup>This relationship between vaso-dilation and rate of pulse has been noted by Bonser, F. G. *A Study of the Relations between Mental Activity and the Circulation of the Blood.* Psych. Rev. 1903. V. p. 125.



records of the last two reactions, but it was difficult to detect in the remaining three, although one of these reactions was described as being "just about all in auditory motor verbal images," which, however, were not distinct to the subject.

The volume curves of this subject showed several types of variability. The volume was sustained evenly in the solving of one reaction. In one reaction, it increased gradually as in the case of Subjects A and B. In another reaction, it showed a constriction following the presentation of the problem,<sup>1</sup> rising into a dilation by the time the solution was reached. The remaining reaction showed a constriction following the presentation of the problem, then a dilation, and finally a constriction at the time when the problem was solved.<sup>2</sup> No noticeable differences in the rate of heart beat or in the height of the pulse curve were found to vary with these changes in volume.

The respiration of Subject C differed slightly from that of Subjects A and B in that she tended to hold her breath during the presentation of the problem and immediately following. Then her breathing became more rapid and superficial as was the case with both Subjects A and B.

Subject D solved two problems during both of which "vocal motor imagery" was reported present. Both records showed noticeable inner speech with occasional laryngeal activity. The amount of inner speech, however, did not account for the verbal imagery which she described. It appeared at the beginning of her thinking and disappeared very soon after, in spite of the fact that she reported the presence of verbal imagery throughout. This subject was the one who reported that she experienced no conflict between her verbal imagery and the verbal distractions of Sitting IV.

The plethysmographic curve in one reaction showed a slight constriction shortly after the presentation of the problem. This was accompanied by noticeable hand movements and was followed by a very gradual increase in volume. The height of the pulse curve was slightly diminished as the volume increased. Toward the end of the reaction a feeling of self-consciousness was reported by the subject. The volume

<sup>1</sup>Such a fall in the volume curve has been reported by Binet, A., and Courtier, J., *op. cit.* p. 162, by Angell, J. R., and Thompson, H. B., *op. cit.* p. 34, 39, by Shepard, J. F., *Organic Changes and Feeling*, *Am. J. of Psych.* 1906, XVII. p. 580, and by Stevens, H. C., *op. cit.* p. 468.

<sup>2</sup>*Cf.* Angell, J. R., and Thompson, H. B., *op. cit.* p. 50.

curve showed a slight constriction, an effect opposite to that found when Subject A was experiencing the same emotion.<sup>1</sup> The other reaction showed no volume changes, no change in height of pulse curve or rate of heart beat, and but one instance of hand movement during the first part of the thinking when she reported she was slightly distracted by the feeling of her pulse.

The respiration of Subject D was similar to that of Subject C in that she tended to hold her breath during the presentation of the problem, but it differed from that of any of the other subjects as the breathing remained superficial up to the last.

A distribution of the different kinds of movement recorded in this sitting is given in the following table, Table XV, throughout five-second periods.

From these results it is apparent that tongue movements were not as frequent in the later parts of instances of prolonged thinking. Laryngeal movements accompanied only part of the tongue movements. Both the rate of breathing and of heart beat were noticeably regular, despite the length of reaction time.<sup>2</sup> However, the rate of heart beat showed a slight tendency to alternately decrease and increase throughout the thinking. The rate of breathing, on the other hand, showed a slight tendency to increase and then to decrease until the problems were solved.

As the total time given to thinking in this sitting was 513.3 seconds, eliminating the one reaction in which the tongue tambour failed to record, the amount of time during which tongue movements were present included 34.7 per cent of the total time of thinking. Vertical movements of the larynx accompanied 22.4 per cent of the tongue movements of inner speech. The horizontal movements of the larynx were less frequently corded, accompanying only 12.4 per cent of the tongue movements.

The table, Table XVI, gives the amount of each kind of movement found in each quarter of the thinking in terms of the percentage of the total time for each quarter.

---

<sup>1</sup>Emotions show vaso-constrictions according to Angell, J. R., and Thompson, H. B., *op. cit.* p. 58.

<sup>2</sup>The regularity of the bodily activity accompanying heightened attention has been noted by Binet, A. and Courtier. J. *op. cit.* pp. 158 ff. and by Angell, J. R. and Thompson, H. B. *op. cit.* p. 67.

TABLE XV.

5 sec. Periods	No. Problems	Av. time Tongue M'v'ts.	Av. time Vert. Lar. M'v'ts.	Av. time Hor. Lar. M'v'ts.	Av. No. Heart Beats	Av. No. Respirations
1	12	2.5 sec.	.30 sec.	.46 sec.	6.4	1.4
2	12	2.5	.49	.45	6.4	1.4
3	11	2.1	.89	.24	6.4	1.4
4	10	1.6	.53	.22	6.7	1.4
5	8	1.02	.37	.12	6.6	1.6
6	8	.8	.08		5.3	1.3
7	6	.8	.20	.04	5.7	1.1
8	5	1.8	.75	.50	5.7	1.4
9	4	2.3	.25	.05	7.1	1.8
10	4	1.	.25		7.3	1.6
11	4	2.2			7.2	1.6
12	4	.2			5.6	1.1
13	3	1.	.06	.06	7.5	1.
14	3	1.	.33		5.5	1.
15	3	.6			5.8	.9
16	2	1.	1.		8.2	1.3
17	1				6.5	1.
18	1	1.		.50	6.5	1.
19	1				6.5	1.
20	1	.5		.50	6.5	1.
21	1				6.	1.

These results show that inner speech tended to occur more frequently at the beginning of the thinking, being seldom apparent during the last quarter of the reaction. The verbal el-

TABLE XVI.

Quarters	Tongue M'v'ts.	M'v'ts. of Vert. Lar.	M'v'ts. of Hor. Lar.	No. of Heart Beats	No. of Respirations
I.	51.5%	6.4%	7.8%	178.5	33.9
II.	28.1	9.8	2.8	180.5	39.4
III.	35.1	5.1	2.6	176.2	40.4
IV.	13.7	7.4	2.6	181.2	39.
Av.	32.1	7.1	3.9	179.1	38.1

ements which appeared in the last quarter, however, were more apt to involve both tongue and laryngeal movements than were those in the earlier part of the thinking. The pulse rate showed a slight alternating increase and decrease,<sup>1</sup> reaching its height in rate at the end.<sup>2</sup> This seemed to be independent of the changes in rate of respiration as recorded in the last column of the table.<sup>3</sup> The breathing tended to increase

<sup>1</sup>Similar alternations of pulse rate have been reported by MacDougall, R. op. cit. p. 163, and by Angell, J. R. and Thompson, H. B. op. cit. p. 68.

<sup>2</sup>Cf. Binet, A. and Courtier, J. op. cit. p. 163.

<sup>3</sup>Cf. Binet, A. and Courtier, J. op. cit. p. 139 and 158, and Angell, J. R. and Thompson, H. B. op. cit. p. 68.



after the presentation of the problem, and then to decrease slightly.

Of the twelve reactions in this sitting, seven showed the presence of the initial pause during which no inner speech could be detected in lip or laryngeal movements. The average length of this pause was 2.1 seconds. Five reactions only showed a final pause, just preceding the subject's final signal. The average length of this pause was longer, 5.3 seconds. This decrease in the number of initial and final pauses reported in Sitzings I-III may indicate that minute speech movements were being made by the subjects during these pauses which could not be observed by the experimenter. But it is noteworthy that these pauses did exist even when the apparatus was delicate enough to detect the slightest articulatory movements.

An examination of the subjects' retrospections of the reactions which showed these final pauses, discloses the fact that in each case verbal elements were present in consciousness throughout the novel part of the thinking, although these verbal elements were not noticeable in the record of tongue and laryngeal movements. Subject A described her verbal images as faint and said that she herself could not tell whether or not they were kinaesthetic. Subject C reported that she could perceive motor accompaniments with her verbal elements during the early part of her thinking, but she did not refer to the nature of the verbal elements which followed. In the other three instances these verbal elements were described as "auditory-motor," "vocal-motor," and "auditory-kinaesthetic" verbal "images," which they undoubtedly were, as they did not involve muscular reactions which could be detected by the most delicate apparatus.

Furthermore, a closer comparison of the subjects' retrospections with the kymographic records shows a surprising correspondence between articulatory movements and verbal elements reported to have been present in consciousness. On the other hand, there is the equally surprising fact that frequently the verbal elements which were described by the subject had no counterpart in movement of tongue or larynx as shown by the kymographic records. This lack of evidence of motor accompaniments for verbal elements in consciousness stands in support of many of the results given

throughout this investigation: that there are verbal images which are not instances of inner speech.

As in this sitting the time during which inner speech was present was recorded, it was possible to correlate speed of thinking with inner speech. As a result,  $r = .606$  which seems to grant to inner speech an important place in thought.

Finally, from the results of this sitting the following conclusions seem justifiable:

I. Many of the verbal elements in thinking are words actually articulated by the individual. Some of these verbal elements involve movements of both the larynx and the tongue. Others consist in merely minute tongue movements.

II. When an individual is uncertain as to the nature of his verbal elements, or even believes them to be images, the presence of accompanying motor reactions may still be detected by the use of delicate instruments.

III. Some verbal elements in thinking, however, are not found to involve articulatory movements, and so are apparently actual images.

IV. The greater part of thinking is carried on in terms other than inner speech.

V. Inner speech, however, plays an important part in thought as the correlation between the amount of the former, as compared with the amount of the latter, is fairly high ( $r: .606$ ).

VI. The vaso-motor and respiratory accompaniments of thinking show a marked regularity, with slight changes brought about gradually. Respiration is more readily affected by thinking than is the volume of the arm or the pulse rate. The respiratory and the vaso-motor functions, moreover, seemed to be independent of each other.

## CHAPTER VIII.

### INDIVIDUAL DIFFERENCES.

Before summarizing the results of this investigation, a consideration of the individual subjects and of their characteristic modes of silent thinking should be given.

#### *Subject A.*

A rapid thinker: all fifteen reactions were instances of rapid thinking. A.D. from the average reaction time for thinking, —22.1 seconds.

An "active" thinker: A.D. from the average in perceptible movements, +25.3 per cent. Minute movements were detected with verbal, visual, and hand-motor images.

Not very dependent upon imagery: A.D. from the average, —20.6 per cent.

Actual type of imagery, visual.

Predominating kind of imagery in novel portions, verbal (inner speech?).

Concrete imagery brought to her the meaning of words, symbolized whole situations, and particularized general and abstract thinking. Several times it appeared and she followed it to novel ideas.

In "real thinking" she said words were present. But five problems were solved without the consciousness of either concrete images or verbal elements, and six of eighteen reactions showed that the novel part of the thinking was carried on in this imageless state. Such conclusions came as "facts" or she was aware of them before they were hardly conscious. Moreover, she reported she once gave up trying to think "rationally" and gave way to the "associative processes" which produced a conclusion which could not be described as verbal or concrete.

After having heard the problem she did not repeat it verbally, unless her attention wandered or she found difficulty in solving it. Usually she retained an "Einstellung" toward the problem which she could not describe in terms of concrete or verbal elements.



*Subject B.*

A moderate thinker in regard to speed: Seven of fifteen reactions were instances of moderate thinking. A.D. from average, +17.1 seconds.

Apparently a quiet thinker: A.D. from average in perceptible movements —20.4 per cent. But minute movements were detected with her visual, verbal, and hand-motor images.

Very dependent upon imagery in her thinking: A.D. from average +35.6 per cent.

Actual type of imagery: mixed, visual and verbal.

Predominating kind of imagery in novel portions, verbal (inner speech?).

Concrete imagery alone was present in one reaction and accompanied verbal elements in eighteen. They apparently came as alternatives which she either accepted or rejected, frequently in verbal terms, depending upon whether or not they were "significant to the train of thought." Once, when no alternative came, she "made a strong attempt to get some kind of imagery." Verbal elements frequently accompanied the concrete images which were "more or less objective. The verbal is personal; no difficulty in having them simultaneously. One is the background of the other."

But these verbal elements were not the act of thinking. The words of the problem were often repeated, accompanied by "a feeling of blankness" in the thought. "I had a long feeling of dullness, with verbal images, but didn't feel I was getting anywhere." When progress was made, on the other hand, the verbal elements were "not whole phrases. Only the high spots are touched. They are fragmentary. I could not give a complete description of them. One phrase lingers until I strike the next."

"Thinking itself goes on in an undercurrent, not exactly verbal. I am conscious of a phrase which I then express. Then I go on to the next." "A running comment keeps going on. It is hard to detect. There may be pauses but it is hard to detect them." "Thinking seems also to imply inflections and modulations of the voice, auditory and motor imagery, which is not verbal." A sudden solution she described as "a sort of quick catching at an idea" which she found to be in verbal form. Her inner speech, as shown in her records of Sitting V, was only occasionally present, showing spaces in which this

"running comment" may have been present, but not perceptible as speech.

When the problem was presented, she usually repeated it in verbal terms when certain words were emphasized. "The problem may go on repeating itself with different emphasis on different parts." Sometimes this became mechanical: "After several moments of mechanical repetition of the problem, I made myself get an impetus to go on." But usually this repetition is resorted to merely until the "problem becomes synthesized. Then I have a sudden feeling that I have the meaning compactly. After the meaning is clear, "there is a settling down to the problem." Then comes "waiting for the set to bring up something."

*Subject C.*

Average in speed: Eight of fifteen reactions were instances of rapid thinking. A.D. from the average, —.5 seconds.

Apparently a quiet thinker, A.D. from average in perceptible movements, —27.4 per cent, but minute movements were detected with her visual and hand-motor images and with part only of her verbal elements.

Not very dependent upon imagery, A.D. from average, —29.3 per cent.

Actual type of imagery: mixed, visual and verbal.

Predominating kind of imagery in novel portions, verbal.

Concrete imagery appeared with the verbal elements in all of her reactions, but six of the twenty-four actual solutions were presented in terms of concrete imagery independent of the verbal. Imagery seemed to present alternatives, as in the case of Subject B. When she was "casting about" for alternatives there was no imagery, merely feelings of strain sensations. Without imagery she felt "receptive" having no feeling of progress. But when the images appeared, concrete or verbal, she "commented" on them, accepting or rejecting them in verbal terms.

The characteristics of this commentary "were hard to detect": "When I talk to myself I haven't a good idea when images start or stop. These images are so familiar they don't catch my attention." As so little inner speech was perceptible in her records of Sitting V, as compared with her retrospections, her thinking may have been carried on

in terms of verbal images, with the exception of the blank periods which she described. However, very little imagery was reported by her at any time.

When the problem was presented, she seldom repeated it in verbal terms, as it usually called to her mind visual images. "When I hear a problem, one or more words are emphasized so that I can find associations for them."

### *Subject D.*

A moderate thinker: Seven of fifteen reactions were rapid, and seven were moderate. A.D. from the average, +11.4 seconds.

An active thinker: A.D. from average in perceptible movements, +14.7 per cent; also minute movements were detected in her visual and hand-motor images and also in her verbal images, although she said these could be carried on independently of vocal movements, even during verbal distraction.

Very dependent upon imagery, A.D. from average +20.3 per cent.

Actual type of imagery: mixed, visual and verbal.

Predominating type in the novel part of her thinking, verbal with visual.

Concrete imagery seemed to have at times the important place in her thinking: "Visual imagery was present all the time, the verbal now and then" or "verbal images were subordinated to the visual. They lacked detail." On the other hand, many times the verbal images were the more important. "The visual images were not important." "Visual images were very fleeting. They had very little to do with the process." "I think the verbal preceded the visual. The visual lingers longer and starts trains of ideas." When they appeared simultaneously, "I would pay attention to one, then to the other" or "sometimes one may be in the foreground, sometimes the other."

She, too, talked to herself a great deal, "While I was seeing this I was saying words here and there to myself." "I said to myself what I would do." Moreover, the following mental states were described at times as verbal: "I wondered," "I had difficulty in deciding" and "I thought." She frequently reported feelings of blankness, especially after she had thought of a possibility and rejected it. And once, with a



"gap" in her flow of imagery, she reported: "I don't feel that progress has stopped."

She seldom repeated a problem in verbal terms, as it usually brought immediately to her mind visual images. But with certain problems, she repeated the words "several times," frequently repeating only part of them.

### *Subject E.*

Average in speed: Seven reactions were rapid and six moderate. A.D. from average,  $+.2$  seconds.

An active thinker: A.D. from average in perceptible movements,  $+27.3$  per cent.

Minute movements were found to have accompanied her visual, verbal, and hand-motor images.

Very dependent upon imagery, A.D. from average,  $+20.5$  per cent.

Actual type of imagery: mixed, visual and verbal.

Predominating type in the novel part of his thinking: visual with verbal.

Concrete images at times held a prominent place in the thinking; they clarified meaning, and offered specific evidence when needed. "I thought of it very quickly. Then tested it out. I depended upon visual images for the test." "The visual image was the proof. The idea came first and I tried it out." Usually, however, the verbal elements dominated. "I had verbal imagery with a background of visual pictures." "My imagery was mostly verbal, very few visual images."

Her verbal images often represented remarks to herself: "That won't do." "No, you can't do it, you must reason it out." "I asked myself, are they the best ones?" But they did not represent all of the thinking: "I am never sure of words, but I think quite often they are scattered throughout." They were rather used to make thought explicit: "The verbal elements were to emphasize it. I clinched the idea with words." Again they were used in "pinning" herself down: "I use words when I make an effort to attend." "I use words when I am disturbed."

Thinking itself, however, seemed to be in other terms: "The idea came. I don't know the terms of the idea. I don't think it was verbal." Again, "People and names popped into my head. I held them to test. Definite, but I don't know in what terms." "I knew Farrar before I had her name."

"I thought of cream and cereal. I thought of lots of staple products" without consciousness of imagery. Moreover, imagery was not present in the "transitive periods."

Usually the problem called up visual images to mind so that she did not repeat the words. But a few times, "I heard myself say some of the words of the problem, but I felt no movement in my vocal cords," or "I kept hearing words." These seemed to be the means of keeping the problem in mind. In one instance, however, she nodded her head to emphasize the words and fix her attention. If the meaning wasn't clear she became distracted, so that such measures were necessary. And if the solution did not "come" she had vague feelings of "discontent" and "hopelessness."

### *Subject F.*

Average in speed: Seven reactions were instances of rapid thinking while six were of moderate speed. A.D. from average, —2.9 seconds.

A very quiet thinker: A.D. from average in perceptible movements —27.8 per cent.

However, minute movements were detected with his visual and hand-motor images. The verbal distraction caused him annoyance in Sitting IV; but he claimed to have been able to carry on verbal images independently, although he paused in the verbal distraction perceptibly once.

Very dependent upon imagery: A.D. from average, +37.5 per cent.

Actual type of imagery: mixed, visual and verbal.

Predominating type in the novel part of her thinking, verbal.

Concrete imagery seldom played a nimportant part in his thinking. A few times the situation was all "visualized," but verbal elements were also present. Once the solution, a visual imagination image, preceded the verbal elements. Usually, however, his concrete images only "helped a little." "Visual image doesn't add. It gives no confidence. Not clear. The images are symbolic. Ideas are distinct and come first."

His verbal elements were usually remarks to himself. "I began to talk to myself." "I think I was talking to myself." "I found myself saying to myself," etc.

His thinking, however, seemed to be something else: "Verbal imagery checked me up, but didn't solve it. It keeps

you on the track." One conclusion which he found hard to describe was "fitted into old principles. Seemed to fit. Gave feeling that it fitted." Ideas also were "not imagery necessarily." However, his reactions showed that the verbal elements were present throughout most of his thinking.

He often repeated the problem once to himself: "I had to repeat it. I had to sort of force myself to branch off and get started on it." "I started to repeat the question as soon as I heard it. It helped fix it in my mind. It helps solve the question. It helps me get started. Certain things come up." And again, "Images came up in answer to a felt need."

### *Subject G.*

A rapid thinker: Thirteen of fifteen reactions were rapid: A.D. from average, —22.5 seconds.

Average in his amount of perceptible movement: A.D. from average —8 per cent.

Minute movements were found to accompany his visual, hand-motor, and verbal elements.

Not quite as dependent upon imagery as the average, A.D. —11.1 per cent.

Actual type of imagery, visual.

Type predominant in the novel part of his thinking, verbal, with an equally large number of states impossible to describe in concrete or verbal terms.

Usually his visual images were "vague" and "fleeting." "A tiny scrap of imagery seems to stand for a lot of meaning." An image of a submarine was "schematic"; of an automobile, "generalized or symbolic." Once, however, he apparently had a clear image. "I had a picture of an angle outside there," pointing. Later it had lines marked off in it, but then he "was not sure how they appeared."

His verbal images were also hard to describe. From the first, he said he didn't know what a motor image was, so he described his verbal images as auditory. Again, he reported they might be motor and once he admitted: "If I did anything, I said it to myself." In the experiment with the verbal distraction, he gave up one problem, saying "I felt I could solve the problem, if I could talk." Moreover, during the solution of the problems he frequently reported that he had the desire to speak out loud. But these verbal elements were not distinct. When he described a novel conclusion as verbal he was always



in doubt about it. "I wasn't conscious of saying this in detail. It was vague." "Verbal, if anything." "I suppose it was verbal."

There were a great many occasions when he could not describe his thinking. Monroe's book came to him as a "thought." "I don't know how I thought of water being staked out." "I had a feeling of relation." Again, "it was all mixed up like a jumble. I couldn't seem to make it clear. It became clear finally." Usually the novel conclusion "all came at once," or "it came quickly with no introduction." Once, however, "the solution came in steps. Sometimes I felt as if I wasn't thinking." The larger part of his thinking was in this indescribable form.

When the problem was presented, it frequently brought up visual images which started his thinking. Usually, however, he had auditory verbal images of one or more words. They "stuck in my mind," "they seemed to stay in mind. They seemed to reverberate and help me stick to the problem." The problem also aroused "sets" in his mind. "I had two sets; one of seeing the city, the other, in an hour"; "a half a dozen sets came before the problem was stated."

### *Subject H.*

A rapid thinker: All reactions were instances of rapid thinking. A.D. from average, —27.4 seconds.

An active thinker: A. D. from average in perceptible movements, +23 per cent.

Minute movements were found to accompany his verbal elements, but not his visual images.

Average in the amount of imagery reported: A.D. from average, +9.3 per cent.

Actual type of imagery: mixed, visual and verbal.

Kind of imagery predominating in novel part of thinking; visual.

Concrete imagery played an important part in his thinking. It was present in every reaction. Alone it solved two problems and represented seven novel conclusions. It was frequently brought to mind by the presentation of the problem, and in one instance the problem was "rethought" in purely visual terms. These images were usually "fleeting."

Verbal imagery was also frequent, once with visual images

"Flickering on top like snapshots." He often questioned himself: "which are the two best" etc.?

When the problem was presented, he frequently reported "auditory verbal images as usual." He recalled the words as they sounded, only occasionally repeating them himself in motor form. "In awareness of problem, imagery was used in details. Sometimes clear and distinct. But in grasp of problem, I was not aware of imagery." When the whole grasp of the problem was not present, the problem was repeated: "When the problem was stated, I had a sort of accumulative idea of the whole. I sort of held it up to the end." "I had a general composite idea of what the problem meant. I won't deny imagery, but it was not definite. Imagery followed immediately. Then I know what I am going to do and what the problem is: a kind of a feeling which carries force."

A good deal of his thinking, however, he could not describe in verbal or concrete terms. "The solution comes in a flash. I don't know what its constituents are: a bare fact, not having specific imagery. It won't persist, but it is more clear than anything else." "Solution comes like a flash. Recognized as such. I might later test it out. Usually I think there is no need of thinking any more about it, and so I turn to the next thing." "Imagery was nearly always present except when I was suddenly realizing the solution or grasping the problem. A sudden realization of the thing I am seeking. It comes as a whole." "I sort of realized that the situation was a composite affair. I didn't think out but a few consequences. I felt I could figure out a good many if I took time." On the other hand, he once experienced a "moment of strain following the problem. Felt expectant until the images followed. I don't know what is present in those periods. I open up the gates to see if something won't come in. If nothing happens, distraction follows and mind wanders." "There is no progress in periods of strain, but they afford opportunities to progress. They are necessary to progress. These periods of strain are periods of the hardest work of all."

#### *Subject I.*

A slow thinker: Eight of fifteen reactions were slow. A.D. from average, +48.6 seconds.

Average in amount of perceptible movements: A.D. —3.5 per cent.

Minute movements were found to have accompanied his verbal and hand-motor imagery but none in connection with his visual images.

Dependent very little upon imagery, A.D. from average —47 per cent.

Actual type of imagery: mixed, visual and verbal.

Kind of imagery predominating in the novel part of thinking, visual with verbal.

Usually he described his visual images as visual "awarenesses." "Visual awareness is not imagery. A difference of degree, they are not like actual things. I have not had many visual images. Sometimes these seem real, but usually they are indistinct. They may be of three dimensions. I can think of vague visual awarenesses and of vivid visual awarenesses, but they seem different from images." Once he reported, "when you were reading the problem, I came as near a visual image as I can get." These visual "awarenesses" were "very weak," "not very clear," "fleeting." They played an important part, however, for they "presented suggestions" and "images were present when I was succeeding."

His verbal elements he recognized oftentimes as inner speech: "auditory-motor verbal images which may have been sensory"; "there was actual speaking"; "at times I do talk to myself as I go along. Sometimes I have words as if spoken out to you." These elements were apparently used especially to give definiteness and clarity to his thinking. When the bell was sounding in Sitting IV, he several times spoke out loud. Again, "the problem seemed very abstract and ideal. I used verbal imagery more for that reason." "Inner speech repeated the catch words, "aesthetic," "ideal," etc.

"Thought," he said, "if not inner speech, is kinaesthetic, more actually verbal." The movements, then, seemed to him to constitute the thinking, for kinaesthesia played a frequent part in his thinking. "Acceptance is probably kinaesthetic." "Identifying myself with the person was kinaesthetic." With his idea of generosity he had a "kinaesthetic image," and the feeling that accompanied the "rich apartment" was *high* but of the "poor relations" *flat*. There were blank feelings during his thinking: "Blank periods make you feel you



aren't gaining;" but finally alternatives or suggestions "come up" and the thinking progresses. His consciousness of the term "tablets" he described once as "like an imageless thought." Aside from this instance which he could not otherwise describe, he reported the presence of three solutions which he could not account for. They came so fast he could not discover what they were, but he had the feeling that they "fitted in." They seemed to be only "general ideas."

The problem usually called to mind visual images, so that he did not often repeat the question. He received from the start "a sort of set."

### *Subject J*

A rapid thinker: Twelve of fifteen reactions were instances of rapid thinking. A.D. from average, —11.8 seconds.

Average in his amount of perceptible movements: A.D. —2.5 per cent. Minute eye movements were found to have accompanied his visual images.

Not very dependent upon imagery in his thinking. A.D. from average, —16 per cent.

Actual type of imagery: visual.

The majority of his novel conclusions he could describe neither as concrete nor verbal.

Visual imagery played a very important part in his thinking. In sixteen of the nineteen reactions it was the only imagery reported. It was the only kind of novel solution which he could describe. It was usually "vague," "scrappy," "schematic," "faded away," "not distinct or thick." Sometimes it was "not important," and "insignificant," but usually it symbolized a situation: a "map with lines in the direction of England," a "map" with "dots on the map of the submarine section," or a "sweeping schematic view of South America." These schematic images he termed "notions," although he did at times differentiate this term from the image: "not clear, no image, just a notion."

Distinct verbal elements played very little part in his thinking, as they were reported in but three of his reactions. These were "auditory-verbal, no motor, and repeated the words of the problem or one or two words in the thought." In the experiment with the verbal distraction, the problem was readily solved with the consciousness of only a few visual images, and no feeling of distraction was present.

There were, however, three short pauses in his repetition of the letters. They may have been breathing pauses, but the reaction was merely 16.6 seconds long, which shows some inhibition may have been present. He also spoke out loud, during the visual distraction.

His conclusions came to him usually "like a flash." Of one he said: "not visual, I can't make out any sensory elements, it just came." Of another he said: "No terms at all. In terms of itself. Distinctly non-sensory. I was aware of it and it was clear," adding: "I am as sure of what I have told you about my introspection, as I am about after-images." He also reported feelings of "direction," "determination," and "tendency forward," as well as "abstract" thinking.

When the problem was presented it usually brought to mind visual images. "The first thing I must do is to get the problem set. I would get along better than to go ahead without knowing what the problem is (not verbal, bare thought)"; "throughout is a feeling of determination. What I am doing is in harmony with the task: no mind wandering." Whether this set was usually present or not, he did not report.

He seldom mentioned talking to himself, but many of his reports were of this sort: "I thought, how else could you do that?" "Then I asked myself, can we?" "Then I thought, oh yes, rote memory," etc. "Then I thought there is no need of staying here, for it is not new." "Then I thought I am no better off here than before." These instances were not described as verbal elements or recognized as such, but they seemed to represent a discursive element which characterized his thinking. It could not have been detailed or distinct, for it was sharply differentiate from the verbal elements; but that form of report was given so constantly it seemed to have been typical of his thought processes.

## CHAPTER IX.

### GENERAL CONCLUSIONS

Silent thinking, or the reaching of novel conclusions without the aid of explicit forms of behavior, may be carried on in very short periods of time.<sup>1</sup> It may be accompanied by a consciousness of concrete imagery<sup>2</sup> without any consciousness of verbal elements.<sup>3</sup> Or, it may be accompanied by a consciousness of verbal elements independent of concrete imagery, as when the individual talks or argues to himself.<sup>4</sup> These extremes, however, are not frequent, for thinking is usually accompanied by a consciousness of both concrete imagery and verbal elements.<sup>5</sup> In general, the more rapid the thinking, the more frequent are these details in consciousness.<sup>6</sup>

This consciousness of concrete and verbal details accompanies but part of thinking.<sup>7</sup> The remaining part of thought is imageless consciousness, independent of both concrete imagery and verbal elements. This imageless consciousness intervenes between periods of imagery.<sup>8</sup> Some of it may be accounted for in terms of inhibition in mental processes.<sup>9</sup> But a greater part of it bears an intimate relation to the development of the thought.<sup>10</sup> In fact, this imageless consciousness frequently characterized the actually novel part of the thinking,<sup>11</sup> and even accounted for the entire solution of some instances of rapid thinking.<sup>12</sup>

An individual is noticeably quiet when he is actually thinking, for a pause in movement is usually apparent just before

---

<sup>1</sup>Average time of 150 reactions was 50.7 seconds.

<sup>2</sup>Problems were selected which would stimulate the presence of concrete imagery as much as possible.

<sup>3</sup>In 24 of 150 reactions.

<sup>4</sup>In 11 of 150 reactions.

<sup>5</sup>In 110 of 150 reactions.

<sup>6</sup>In Sitting II, speed of thinking and imagery were correlated .45 while rapid thinking and imagery were correlated .55. In Sitting III, these correlations were respectively .21 and .59. Of the combined 100 reactions, these correlations were respectively .18 and .506.

<sup>7</sup>51.6% of time in 100 reactions.

<sup>8</sup>In 99 of 150 reactions.

<sup>9</sup>89 of 252 such periods were described by the subjects as "breaks," "gaps," and "periods of waiting."

<sup>10</sup>130 of 252 such periods were described by the subjects as being accompanied by a feeling of progress.

<sup>11</sup>In 21 of 150 solutions.

<sup>12</sup>In 5 reactions.



the thinker indicates that he has reached a novel conclusion.<sup>1</sup> Such a pause is noticeable, even when the implicit movements of inner speech are being measured.<sup>2</sup>

Throughout thinking, slight vaso-motor and respiratory changes take place. The volume of the arm shows a slight tendency to increase gradually until the novel conclusion is reached.<sup>3</sup> The rate of heartbeat shows a slight increase at the beginning of the thinking and then alternately increases and decreases until the conclusion is reached.<sup>4</sup> Respiration tends to be more superficial and rapid after the problem is presented, becoming more regular as the solution is being reached.<sup>5</sup>

Verbal elements in thinking are frequently accompanied by noticeable lip movements<sup>6</sup> and by minute movements of the tongue and larynx.<sup>7</sup> Visual images in thinking may be closely related to ocular movements<sup>8</sup> and in many instances are inseparably connected with them or with a condition of visual fixation.<sup>9</sup>

The failure of imagery to be constantly present in thought, the lack of verbal elements in the solving of many problems, and the absence of both in some instances, point to thinking as a conscious process underlying these varying accompaniments.

Although implicit forms of behavior accompany thinking to a noticeable extent,<sup>10</sup> even these movements tend to disappear as the novel conclusion is being reached. Thinking takes place too rapidly to consist merely in a series of sensori-motor reactions. The time taken in reaching a new conclusion is very much less than the time needed to describe the same process to others. With memory visual images, slight eye movements might accompany the reproduction through associa-

---

<sup>1</sup>In 115 of 150 reactions. This pause was not as conspicuous in solutions which were discarded because they were not novel to the subject (21 of 49 reactions showed this pause).

<sup>2</sup>In 5 of 12 reactions.

<sup>3</sup>In 7 of 12 reactions.

<sup>4</sup>In 6 of 12 reactions. The remaining pulse records show no common characteristics.

<sup>5</sup>In 6 of 12 reactions. The remaining breathing curves show no common characteristics.

<sup>6</sup>Observed by the experimenter in 32 of 74 instances.

<sup>7</sup>See above Table XV.

<sup>8</sup>Observed by the experimenter in 37 of 92 instances.

<sup>9</sup>See above pp. 59 ff.

<sup>10</sup>Minute relevant movements were those most apparent in Sittings I and II., while the correlations between thinking and inner speech in Sitting V was .606.

tions with the visual details originally seen, but these would not be necessary for visual imagination images.<sup>1</sup> Familiar verbal phrases might through habit reproduce their corresponding articulatory movements. With novel ideas, on the other hand, such movements would not be as likely to occur.<sup>2</sup> Moreover, movement is sometimes intentionally resorted to in order to make details clearer, since it arouses sensations of the movements, which may in turn arouse further associations.<sup>3</sup>

Furthermore, the arousal of movement by imagery requires a strength of stimulus adequate to pass from the sensory areas out through the efferent neurones. Visual and verbal images, which have been differentiated from others of their kind by merely their lack of clearness and detail, may be examples of imagery so weakly stimulated that the energy from that center is not sufficient to arouse accompanying movement.<sup>4</sup> This seems to afford evidence that is in support of the theory that thinking may be other than a mere series of sensori-motor arcs.

Thinking is a matter of association and, as this is a physiological process, it takes care of itself largely.<sup>5</sup> When the problem is presented the individual perceives the words and their meaning, but is at the same time aware that they constitute a problem. They describe a situation which is novel or unfamiliar, to which he must react. The novelty lies in the relationship which the terms of the problem bear to each other. The individual becomes aware of this relationship and realizes that he cannot solve that problem until he can reproduce the same relationship in terms of his own past experience. Take for example this problem: What picture would you select for the walls of a fifth grade school room? The terms are familiar, the novelty consisting in the relationship of the terms to each other. The individual realizes that the relationship is not familiar and that he has not a duplicate of this in his past experience, so that he cannot at once answer the question. To him, then, this constitutes a problem. By as-

---

<sup>1</sup>37 of 69 memory images were found to have been accompanied with ocular movements while but 17 of 47 imaginary images have this characteristic.

<sup>2</sup>See above pp. 69 ff.

<sup>3</sup>See above pp. 72 ff.

<sup>4</sup>See above pp. 73 ff.

<sup>5</sup>The subjects frequently described the mental processes as progressing independently of their guidance. See above pp. 74 ff.



sociation each auditory sense impression made by the terms in the statement of the problem arouses its associates from the mass of past experiences and the individual attends to this process until some experience, which is partially related to all of the terms, is aroused by the stimuli. In this way the terms of the problem are brought into relationship by the connection of the experience which is related to each through contiguity or similarity. The physiological process is completed, as the stimuli set up by the perception of the various terms of the problem have reached some common association. At this moment the relationship is recognized by the individual as the one desired and the problem is solved. Thinking, therefore, is but a process of the recognition of a novel relationship between elements in one's past experience. The process of trial and error, with its accompanying selection and rejection of material, which takes place in securing the desired relationship, seems to be identical with that to which James has alluded in his description of the recall of a forgotten name finally recognized as the one sought.<sup>1</sup>

The thinker, then, is one who is able to turn his attention persistently to his conscious states. He must be one who can be aware of relationships as such and who can follow mental stimuli as they pass physiologically through brain paths, reproducing past experiences. The relationship required by the elements of the problematic situation must be sought. This is duplicated by a new grouping together of his past experiences, or, when this is lacking, in terms of experiences newly acquired from outside, such as in drawing a diagram, writing an outline, or reading relevant material. The vividness of the conscious states depends upon the degree of motor adaptation in attention, the source of control of the processes of association in thinking, as the degree of adaptation determines the degree of reinforcement of the physical stimuli.

Thinking need not be merely a series of sensori-motor arcs, for much of its activity might readily be limited to the cortical areas of the brain. In this way it might be, to a certain extent, literally a "mental reaction." The stimulation of the various sensory areas at the time the problem is perceived arouse duplicates of past sensory experiences. If stimuli, however, reach motor areas in the brain, and there is sufficient energy for them to be transmitted to the muscles,

<sup>1</sup>James, W. *Psychology*, I. p. 257.



movements will result, the sensations of which arouse in turn stimuli for further associations in the brain. The consciousness attending cortical activity in the so-called "association areas," being neither sensory nor motor, would be accounted for in terms of the active states of imageless consciousness which have been found to characterize so much of silent thinking. As these imageless states occurred not only between periods of imagery i. e. between sensory areas in the brain, but also when novel conclusions were being attained,<sup>1</sup> the association areas might act as centers where stimuli from various areas in the brain can come together. Passive states of imageless consciousness, on the other hand, might be accounted for in terms of inhibitions, blockings in association processes, due to the various paths taken by the different stimuli.

Thinking, as a reaction of this sort, would imply the transmission of the stimuli, aroused by the perception of the elements in a problem, throughout the brain as a nervous center. The stimuli, through associated brain paths, would bring about activity in any part of the brain. The areas stimulated would be of importance only as they proved to be the source of experiences which would make possible the relationship required by the elements of the problem. Hence, images, movements and their consequent sensations, and imageless states of consciousness might, singly or all together, be the content of thought. Such an interpretation of thinking as this best explains the varied results secured in this investigation.

---

<sup>1</sup>Woodworth, R. S. Non-sensory Components of Sense Perception. *J. of Phil. Psych.* etc. 1917, IV. pp 169 ff.

## APPENDIX

### EXPERIMENTER'S CODE

- ☐ Forward motion of head.
- ☐ ☐ ☐ Sidewise motion of head.
- ☐ ☐ Eye movement.
- ☐ ☐ ☐ ☐ Scowl.
- ☐ ☐ Lip movement.
- ☐ ☐ ☐ Smile.
- ☐ ☐ ☐ ☐ Word spoken aloud.
- ☐ ☐ ☐ ☐ Right hand.
- ☐ ☐ ☐ Left hand.
- ☐ ☐ ☐ ☐ ☐ Right foot.
- ☐ ☐ ☐ ☐ Left foot.
- ☐ Movement of body.
- ☐ ☐ Shrugging of shoulders.

## PROBLEMS

## For Sitting I.

1. How would you plan a trip to Iceland?<sup>1</sup>
2. If you were given unlimited funds to make a collection of some particular kind, what choice would you make?
3. What would be the best thing to do if you suddenly found yourself in the midst of an angry crowd which was pushing forward to commit some act of violence?
4. What would you do if you saw a person running in the street after her dress had caught fire?
4. If you were going to be away from all civilization for a year and could take only four books, which ones would you choose?
6. How would you manage if you had a blue serge dress and only half a yard of red silk to trim it with?
7. What do you consider is the best training for a girl (or boy) between fourteen and twenty years of age who is the daughter (or son) of extremely wealthy parents?
8. What do you think is the best way to teach geography to children?
9. During what years is the best time to teach it?
10. If you had only four days for a vacation and \$75 to spend on it, where would you go?
11. Of what would you make your meal if three unexpected visitors came in at lunch time, and you had on hand merely a can of corn and a can of sardines, beside the ordinary supply of staple products?
12. How would you plan a Hall of Fame?
13. If you had to select four pictures to hang on the walls of a fifth grade school room, what would you choose?

---

<sup>1</sup>Suggested by test in Angell, Report of Com. of Am. Psych. Ass. on Standardization of Procedure in Experimental Tests. Psych. Rev. Mon. XIII. 1910. No. 53, p. 98.



## PROBLEMS

## For Sitting II.

1. If you had to give a lecture about the effect of moving pictures on the child mind, how would you arrange your material?
2. Five hundred years ago the automobile, the air ship, and the steam railroad were entirely unknown means of locomotion. If anyone then had said that today we should be traveling by these means, his statement would have been an economic prophecy. Will you yourself make an economic prophecy in regard to future methods of locomotion?<sup>1</sup>
3. Will you make an economic prophecy in regard to future methods of heating?<sup>1</sup>
4. Will you make an economic prophecy in regard to future housing conditions?<sup>1</sup>
5. Will you make an economic prophecy in regard to future eating conditions?<sup>2</sup>
6. How would it be possible for an automobile to be so constructed that it could be used at any moment as a flying machine?
7. How do you think an instrument could be made which would turn the leaves of sheet music?<sup>1</sup>
8. What would be the probable consequences if the seas could be owned by individuals as the land is?<sup>1</sup>
9. If you had only one hour in which to show the city to some visitors what would you do with them?
10. What is the best kind of a sidewalk to build on an up-hill road?
11. What color scheme would you select for decorating the library in your own home?

---

<sup>1</sup>Chassell, L. M. Tests for originality. *J. of Ed. Psych.* VII. 1916, p. 322.

## PROBLEMS

## For Sitting III.

1. What would happen if our supply of paper should cease?<sup>1</sup>
2. If you were alone on a high point of land and had to attract the attention of a person a mile away, what would you do?
3. What would be likely to happen if it were agreed that each individual owned the atmosphere for 500 feet above his property?
4. How would you go about it to find out whether good spelling depended more upon logical memory than upon rote memory?
5. If you were building a house in the country, what kind of a front entrance would you design for it?
6. If the possession of wealth were universally considered dishonorable, how would it affect the cost of living?<sup>1</sup>
7. If you had unlimited funds to found a new periodical, what kind would you choose to start?
8. If you should start now for the Battery, how could you possibly reach there in half an hour?
9. On what historic character could one write a good play?
10. If you were asked to write a magazine article about the Montessori Method, how would you arrange your material?
11. If you were leaving an empty house and your skirt caught in the door which locked itself and for which you had no key, what would you do?
12. How would you entertain a party of blind children?

---

<sup>1</sup>Chassell, L. M. Tests for Originality, *J. of Ed. Psych.* VII. 1916, p. 322.

## PROBLEMS

### For Sitting IV.

1. Visual Distraction (swinging object.)
  1. How would you make an instrument which would rule a blackboard off into squares of any size?
  2. What would be the best kind of a fire-escape to install on crowded dwelling houses?
2. Auditory Distraction (bell rung continually.)
  1. What two musical artists do you think would attract the largest audiences today if they gave a program together?
  2. What do you think is the most inspiring march music?
3. Verbal Distraction (counting aloud a, b, c, a, b, c; etc.)
  1. How would you teach a young child motor co-ordination?
  2. What four points would you especially emphasize if you were giving a college course in literature?
  3. How much money do you think Barnum & Bailey's circus takes in at one performance when they have a full attendance?
4. Motor Distraction (finger play.)
  1. How would you divide an angle into three equal parts?
  2. How would you fold a circular piece of paper so that cutting it in one place, you would have four separate triangles cut out of it?<sup>1</sup>

---

<sup>1</sup>Suggested by Binet & Simon. *A Method of Measuring the Development of the Intelligence of Young Children*. Tr. by C. H. Town, p. 81.



## PROBLEMS

## For Sitting V.

1. If you had only one hour in which to show the city to some visitors, what would you do with them?
2. If the possession of wealth were universally considered dishonorable, how would it affect the cost of living?<sup>1</sup>
3. How do you think an instrument could be made which would turn the leaves of sheet music?<sup>1</sup>
4. If you were asked to write a magazine article about the Montessori Method, how would you arrange your material?
5. If you were to select street signs for some city, what kind would you choose?
6. What occupation is best fitted for the blind to fill?
7. What stories would you tell to amuse a six year old girl for an hour?
8. If you wanted to find the person who owned the oldest piano in the United States, how would you go about it?
9. If a clock's hands were set at quarter of three and the hands were turned back one and two-thirds revolutions, what time would it then be?<sup>2</sup>

---

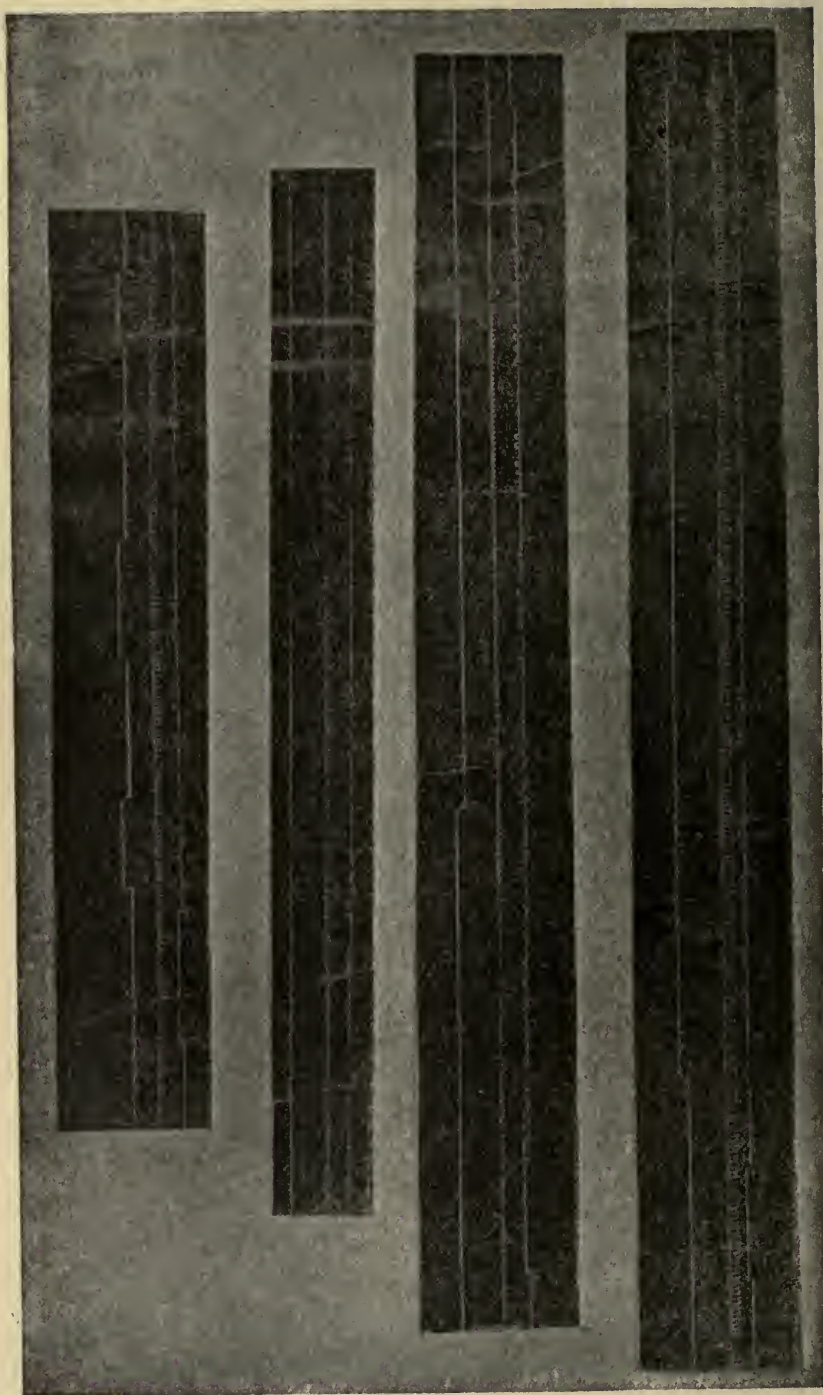
<sup>1</sup>Suggested by test in Chassell, op. cit. p. 323.

<sup>2</sup>Suggested by Test XV. 3, in Goddard, H. H. *The Binet-Simon Measuring Scale for Intelligence Revised*. The Training School, 1911.

## STIMULUS WORDS FOR TEST II IN IMAGERY TYPES

	Problem
1. Dwelling house	IV 2
2. Soldiers	
3. Wealthy parents	I 7
4. Iceland	I 1
5. Dress	I 6
6. Books	I 5
7. Flying machine	II 6
8. Committee	IV III 3
9. Fire	
10. Magazine	III 10
11. Hall of Fame	I 12
12. Battery	III 8
13. Collection	
14. Visitors	I 11, II 9
15. \$1,000	IV, III 3
16. Crowd	I 3
17. Vacation	I 10
18. Accident	
19. Sheet music	IX 7
20. Automobile	II 6

Fig. 8



Typical Kymograph Records of Sitting II.

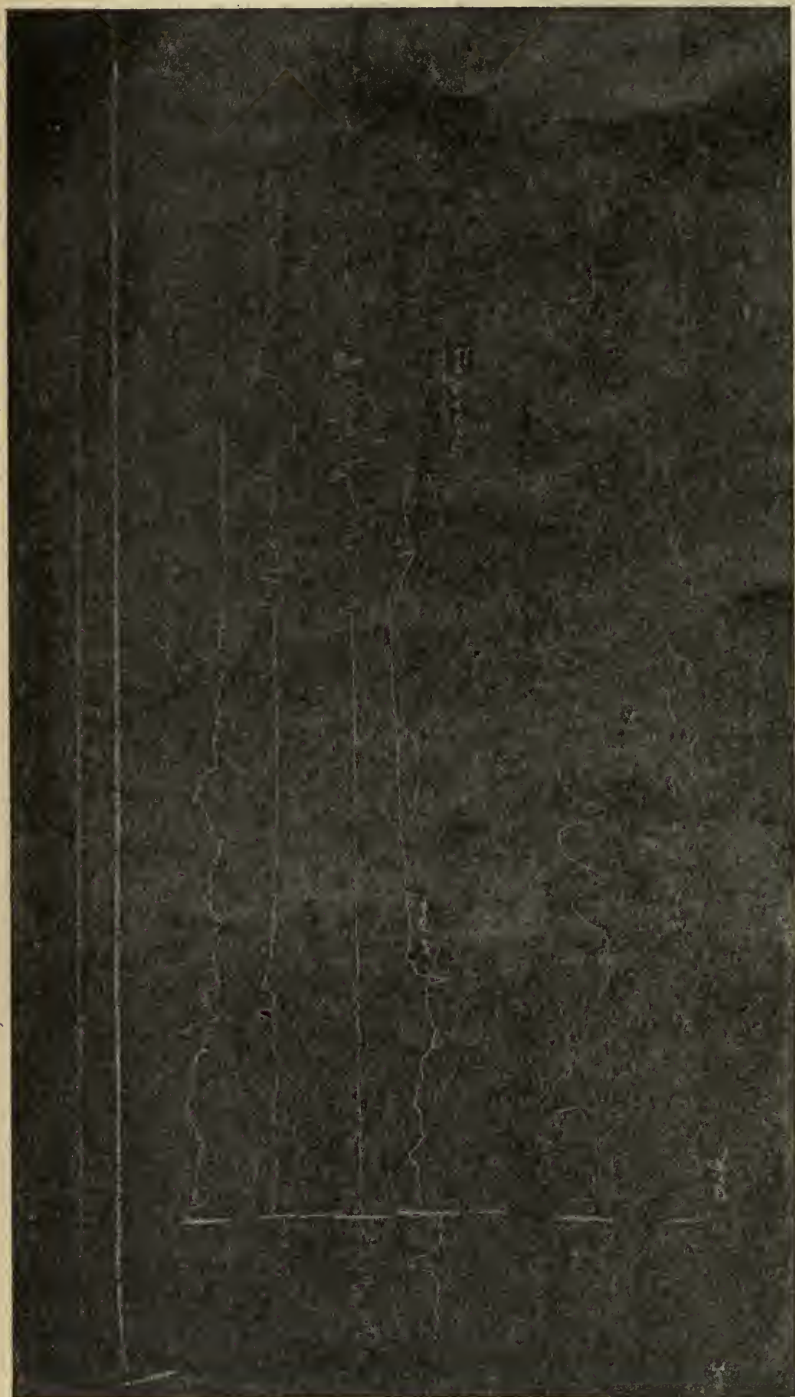
1st line, record of subject's key showing presence and absence of imagery.  
2nd line, record of time in fifths of seconds.





Typical Kymographic Record of Sitting V. Subject A. Problem 8.

- 1st line—Time of presentation of problem
- 2nd line—Time record in fifths of seconds.
- 3rd line—Record of tongue movements.
- 4th line—Vertical movements of larynx.
- 5th line—Horizontal movements of larynx.
- 6th line—Plethysmographic record.
- 7th line—Record of thoracic respiration.
- 8th line—Record of abdominal respiration.



Typical Kymographic Record of Sitting V. Subject B. Problem 9.

- 1st line—Time of presentation of problem and of subject's thinking.
- 2nd line—Time record in fifths of seconds.
- 3rd line—Record of tongue movements.
- 4th line—Vertical movements of larynx.
- 5th line—Horizontal movements of larynx.
- 6th line—Plethysmographic record.
- 7th line—Record of thoracic respiration.
- 8th line—Record of abdominal respiration (after the subject had given the signal that she had finished, she decided that she had given the wrong answer).



Typical Kymographic Record of Sitting V. Subject C. Problem 1.

- 1st line—Time of presentation of problem and of subject's thinking.
- 2nd line—Time record in fifths of seconds.
- 3rd line—Record of tongue movements.
- 4th line—Vertical movements of larynx.
- 5th line—Horizontal movement of larynx.
- 6th line—Plethysmographic record.
- 7th line—Record of thoracic respiration.
- 8th line—Record of abdominal respiration.



## BIBLIOGRAPHY OF REFERENCES

1. Ach, N. Ueber die Willenstätigkeit und das Denken, 1905.
2. Angell, J. R. Methods for the Determination of Mental Imagery. *Psych. Rev. Mon.* 1910, XIII. No 53. pp. 61-107.
3. Angell, J. R., and Thompson, H. B. Organic Processes and Consciousness. *Psych. Rev.* 1899, VI. pp. 32-69.
4. Armstrong, A. C., and Judd, C. H. The Imagery of American Students. *Psych. Rev.* 1894, I. pp. 496-505.
5. Betts, G. H. The distribution and Function of Imagery. Teachers' College, Columbia University. *Contributions to Education*, No. 26, 1909.
6. Binet, A., and Courtier, J. Les Causes d'Erreur dans les Expériences de Plethysmograph. *L'Année Psych.* 1895, II. pp. 88-167.
7. Bonser, F. G. A Study of Relations between Mental Activity and the Circulation of the Blood. *Psych. Rev.* 1903, X. p. 120.
8. Bühler, K. Tatsachen und Probleme zu einer Psychologie des Denkvorgänge. *Arch. f. d. ges. Psych.* 1907, IX. pp. 297-365.
9. Clark, H. Visual Imagery and Attention. *Am. J. of Psych.* 1916, XXVII. pp. 461-492.
10. Cohn, J. Experimentelle Untersuchungen ueber das Zusammenwirken des akustisch-motorischen und des visuellen Gedächtnisses. *Zsch. f. Psych.* 1897, XV. pp. 161-183.
11. Colvin, S. S. Methods of Determining Ideational Types. *Psych. Bull.* 1909, VI. pp. 223-237.
12. Courten, H. C., Involuntary Movements of the Tongue. *Yale Psych. Studies.* 1902, X. pp. 93-96.
13. Curtis, H. S. Automatic Movements of the Larynx. *Am. J. of Psych.* 1900, XI. pp. 237-240.
14. Fernald, M. R. The Diagnosis of Mental Imagery. *Psych. Rev. Mon.* 1912, XIV. No. 58.
15. Feuchtwanger, A. Versuche ueber Vorstellungstypen. *Zsch. f. Psych.* 1911, LXVIII. pp. 161-199.
16. Fox, C. The Conditions Which Arouse Mental Imagery in Thought. *Rep. Brit. Ass. for Adv. of Science.* 1913, LXXXVII. p. 688.
17. French, F. C. Mental Imagery of Students. *Psych. Rev.* 1902, IX. pp. 40-56.
18. Galton, F. *Inquiries into Human Faculty.* London, 1883.
19. Gamble, E. A. M., Attention and Thoracic Breathing. *Am. J. of Psych.* 1905, XVI. pp. 261-292.
20. James, W. *Principles of Psychology*, Vol. I and II.
21. Kakise, W. A. Preliminary Experimental Study of the Conscious Concomitants of Understanding. *Am. J. of Psych.* 1911, XXII. p. 63.
22. Lehmann, A. Die körperlichen Aeusserungen psychischer Zustände, 1899.
23. Lummis, B. A. A study of Thought Processes. *Psych. Studies from Bedford College, Univ. of London.* 1914. pp. 67-161.
24. MacDougall, R. The Physical Characteristics of Attention. *Psych. Rev.* 1898, III. pp. 158-180.
25. Martin, L. J. Die Projektionsmethode u. d. Lokalisation visuelles u. a. Vorstellungsbilder. *Zsch. f. Psych.* 1912, LXI. pp. 321-546.
26. Mayer, A., and Orth, J. Zur qualitativen Untersuchungen der Association. *Zsch. f. Psych.* 1901, XXVI. pp. 1-13.
27. Messer, A. Experimentelle-psychologische Untersuchungen ueber das Denken. *Arch. f. d. ges. Psych.* 1906, VIII. pp. 1-224.
28. Meumann, E. *Uber Oekonomie und Technik des Lernens*, 1903.

29. Perky, C. N. An Experimental Study of Imagination. *Am. J. of Psych.* 1910, XXI. pp. 422-452.
30. Pintner, R. Inner Speech during Silent Reading. *Psych. Rev.* 1913, XX. pp. 129-153.
31. Reed, H. B. Inner Speech in Thought Processes. *J. of Exp. Psych.* 1906, I. pp. 366-376.
32. Seashore, C. E. *Elementary Experiments in Psychology*, Chap. X.
33. Secor, N. B. Visual Reading, A study in Mental Imagery. *Am. J. of Psych.* 1899. XI. pp. 225-236.
34. Segal, J. Ueber den Reproduktionstypus und das Reproduzieren von Vorstellungen. *Arch. f. d. ges. Psych.* 1908, XII. pp. 124-136.
35. Shepard, J. F. Organic Changes and Feeling. *Am. J. of Psych.* 1906, XVII. pp. 522-584.
36. Smith, T. H. On Muscular Memory. *Am. J. of Psych.* 1896, VII. pp. 453-490.
37. Stevens, H. C. A Plethysmographic Study of Attention. *Am. J. of Psych.* 1905, XVI. pp. 409-483.
38. Störring, G. Experimentelle Untersuchungen ueber einfache Schlussprozesse. *Arch. f. d. ges. Psych.* 1908, XI. pp. 1-127.
39. Titchener, E. B. *Experimental Psychology of the Thought Processes*. N. Y. 1909.
40. Washburn, M. F. *Movement and Mental Imagery*, 1914.
41. Woodworth, R. S. Imageless Thought. *J. of Phil. Psych.*, etc. 1906. III. pp. 701-708.
42. Woodworth, R. S. Non-Sensory Components of Sense Perception. *J. of Phil. Psych. etc.* 1907. IV. p. 169.





# SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

BY  
HERBERT WESLEY ROGERS, Ph.D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

No. 49

COLUMBIA UNIVERSITY CONTRIBUTIONS TO PHILOSOPHY AND PSYCHOLOGY  
VOLUME XXVII, No. 4

NEW YORK  
APRIL, 1922

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Cary St., W. C.); Paris (16, rue  
de Condé)

226736  
7. 11. 28  
b



## TABLE OF CONTENTS

1. The Problem.	5
2. History.	6
A. Tests for typists.	6
B. Primitive methods in vocational selection.	7
C. Traditional methods in vocational selection.	8
D. Modern methods in vocational selection.	8
a. Vocational miniature.	
b. Vocational sample.	
c. Vocational analogy.	
d. Trade tests.	
e. Miscellaneous empirical tests.	
3. Procedure.	15
A. Materials employed.	15
B. Individuals serving as subjects.	15
a. Group 1.	
b. Groups 2 and 3.	
c. Age, education and experience of groups 1, 2, and 3.	
C. Methods of obtaining measures of abilities.	17
a. Stenography.	
b. Grammar.	
c. Typewriting.	
I. Group 1.	
II. Description of the work of groups 2 and 3.	
III. Group 2.	
IV. Group 3.	
D. Some remarks on the procedure.	21
E. Administering the tests.	22
a. General procedure.	
b. Instructions to the subjects.	
c. Records.	



4. Results and how obtained.	24
A. Scores in the tests.	24
B. Treatment of errors.	25
C. Methods of combining the measures.	25
D. Correlations.	26
a. Stenography and grammar.	
I. Method used.	
II. Correlations.	
b. Typewriting.	
I. Method used.	
II. Correlations.	
E. Discussion of correlations.	27
a. Significant tests for typewriting and significant tests for stenography and grammar.	
b. Typewriting, group 1.	
c. Typewriting, groups 2 and 3.	
d. Typewriting, groups 1, 2, and 3.	
F. Partial correlations.	29
5. Applicability of the results.	30
Appendix.	41
1. The supervisor as a judge of abilities.	41
2. Averages of the scores in the tests arranged according to age and education.	42
3. Sample of an army trade test.	44
4. Bibliography.	45

## SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

### 1. *The Problem.*

The aim of this investigation is to test out in practice the method of empirical vocational tests and to discover, if possible, significant tests for specific commercial functions.

The type of work selected for correlation with the tests is typewriting. Data in stenography and grammar, obtained with one of the groups of subjects, have been utilized and included in the investigation.

*2. History.**A. Tests for typists.*

Lahy (32)\* reports an experiment made with eleven typists as subjects; six women who had had from two to four years practice, and five men who had had from two to eight years practice, in typing. Measures of the following functions or processes were obtained; muscular sensibility of the two hands, speed of auditory reactions, memory for numbers and sentences, attention, and several higher processes such as imagination, abstraction, judgment and reasoning. In comparing the three good and the three mediocre women subjects it was found that the former possessed a finer threshold of tactile and muscular discrimination, a tendency of the hands to become equally strong, a better memory, fewer errors of attention, but showed longer auditory reaction times. In comparing the men subjects the same distinctions were found. This is an interesting and valuable contribution in that it stimulates work along these lines and blazes the trail, but it has not been worked out extensively enough to be practically applicable.

Lough (35) reports graphical curves showing the relation between the time required to perform a letter substitution test, "habit formation test," measured after repeated trials with the test, and ability in stenography and typewriting for a group of high school students. The curves show a good degree of correspondence between performances in the test and judgments of ability in stenography and typewriting. To quote Lough† "It is evident that if we give these pupils a test in habit formation first, we can determine what pupils are capable of doing good work in typewriting, and we can save some of the pupils a great amount of wear and tear, by simply allowing them to take some other work in place of typewriting. They are not adapted to the motor responses which typewriting calls for. By means of this experiment it would be possible to select those who are likely to succeed in typewriting and to give a vocation to those who would not succeed." If this test could accomplish all that Lough claims for it it would indeed be a boon to the human race but unfortunately, the necessary scientific data concerning the experiment are not given. There is only one test used and the

---

\*Refers to Bibliography.    †p. 94.



investigator who would build up a system of vocational guidance or selection for such a type of work on the ability to perform any one such test, when the original data are inconclusive, would be incautious indeed. However, a good deal of credit is due Lough for his pioneer effort.

Link (33) gives a series of tests to detect ability in type-writing and a series of tests to detect ability in stenography. The tests are, for typing—letter substitution, Trabue completion, context reading, spelling and typing; for stenography—context reading, mixed letters and numbers, substitution, Trabue completion, spelling, grammar, dictation and typing. Link states\* “. . . relevant tests were given to two senior classes of over three hundred girls and boys in a commercial high school, to twenty-six pupils of two business schools, to a group of twenty-two office typists, to another group of nineteen stenographers, to over four hundred candidates for positions as typists and stenographers. . . the tests selected on the basis of these experiments are those which showed the highest and most consistent agreement with the abilities of those examined.” These are broad statements, made in the absence of any norms, methods used, correlations, or any reliable data whatsoever, which should be presented if “consistent agreement” is to be proved.

#### B. Primitive methods in vocational selection.

In primitive magic there is the naive idea that by simply expressing a wish forcibly enough the wish will be materialized; thus if the parents of an Indian boy wished their son to be tall, they would toss him in a blanket. In medieval clairvoyance there was and is still a belief that there are signs which reveal an individual's character. The sources of these signs are principally to be found in the zodiac, palm of the hand, tea leaves, cards, dreams, crystal gazing, etc. In the nineteenth century there was a spot located in the brain for the function of speech and later all parts of the body were found to be localized in the brain. Phrenology was a distortion of this fact. Signs disconnected with the individual's body are now given up and internal and personal external signs take their place. In phrenology there was no comparison of skulls, but individuals were taken at random who had some peculiarity in the formation of their skulls; for instance, if

---

\*p. 422-3 and chapter 8.

an individual had a bump anywhere on his head and was also known to possess some peculiarity such as stinginess, then that bump was labelled stinginess. When it became known that there were no scientific criteria for phrenology, the skull was discarded and the face was taken up, and the pseudo-science of physiognomy was born. Physiognomy maintains that there is something in the features which is significant of inner hidden traits. This is true in so far as the characteristic is an expression of a quality of character. But in vocational psychology we are mainly interested in young people whose characteristics have not been well developed or well stamped upon their faces.

#### C. Traditional methods in vocational selection.

In the selection of employees the employer, for the most part, has been content to form a ready estimate of a candidate's ability by looking him over, or else has accepted him on the authority of someone who has recommended him; or simpler still he tries the candidate at the task and, sink or swim fashion, he promotes him or discharges him according as he succeeds or fails in the work assigned. None of these methods of measuring a man's value will bear criticism. The last is always costly and time consuming; the second is uncertain, depending on the degree of confidence that one has in the recommender, his ability to evaluate and his motives; as to the first there may be some men who are geniuses at picking men of worth at sight; they seem to have an unerring insight which prompts them to their choice, yet even in such cases, it appears that one does not know the signs of value intuitively but the ability to discriminate has been acquired only through long and costly experience.

#### D. Modern methods in vocational selection.

##### a. Vocational miniature.

In the vocational miniature, the work is reproduced on a small scale apparatus which duplicates the actual situation which the worker faces while engaged at his task. Typical of this method is an experiment of McComas of Princeton, in which he attempted to detect good and bad telephone operators. He constructed an actual switchboard on a small scale. The operator, to be tested, made connections at the board which were timed on a kymograph in an adjoining room. This

species of "reaction and co-ordination" time showed the interval between the appearance of a light over a call connection and the moment that the operator "plugged in," and also between the moment that a number was called and the appropriate connection made. Fifty-one trials were made by each of the nine operators tested and the average time of each subject was compared with the average ranking of the subjects by the two supervisors. The test detected the two best and two out of the three poorest operators. The correlation of these rankings with the rankings of the supervisors was .6250 with a probable error of .14. This method has the advantage of concreteness and apparent relevance but there is little actual assurance that the ideas and feelings aroused correspond with the processes aroused in the actual work.

b. Vocational sample.

The method of vocational sampling is closely related to the method of miniatures. It simply consists of taking an actual piece of the work to be performed and sampling the candidate's ability by the degree of his success in the trial. Thus, in testing applicants for jobs as filing clerks, they may be given a number of cards to be arranged in order according to date, number, alphabetical position, address, etc. In certain cases specimens of work have been devised or taken into the psychological laboratory and the worker watched more carefully and measured more exactly. Unfortunately this is a method which cannot be applied to very many kinds of work, for care must be taken to make the tests representative of all the work. At its best this method detects only the presence of ability and is not a means of gauging potential ability.

c. Vocational analogy.

This method consists of an attempt to create, in the laboratory, a situation which arouses and exercises the same mental functions which the actual job arouses and exercises. The material is new, but the attitude and endeavor of the worker are intended to be the same. Munsterberg stated (37) that, by this method, sea captains could be selected who could be relied upon to avoid accidents. The test material consisted of twenty-four cards, the size of playing cards, on which are printed four rows of capital letters namely, A E O U in irregular repetition and interspersed with a few other letters. The person to be tested had to distribute these twenty-four cards as quickly as possible into four piles according



to the numerical preponderance of the four capital letters. The time taken to sort the cards and the number of errors made is taken as the index of ability to safely guide a ship's course. The author of this test admits that there was no correlation scientifically determined between the test and the actual ability to guide a ship's course with safety. The insuperable difficulty with this method is that there is no way of ascertaining that the same mental functions which are called into action on the job are really being exercised and tested in the experiment. This is a highly dangerous method and has done much to put vocational psychology into disrepute in the eyes of men of practical business affairs.

d. trade tests.

The method of trade tests is that of putting specific problems, which are germane to the work to be performed on the job, to the applicant for a job. The answers to the questions give an index of his experience and ability on the job. This idea has been made use of in the past in the examinations of candidates for positions by some concerns and chiefly by some civil service commissions. (Appendix 3 for sample of trade test.)

A good trade test should differentiate between various grades of skill, should produce uniform results in various sections of the country and in the hands of individuals of widely different characteristics and should consume the least amount of time and energy consistent with satisfactory results. While there are all degrees of trade ability among the members of any trade, the method of the trade tests has classified all of the members of any trade into a few groups, usually four, such as novice, apprentice, journeyman, and expert. The novice is a man who has no trade ability whatever, or at least none that could not be paralleled by practically any intelligent man. An apprentice has acquired some of the elements of the trade but is not sufficiently skilled to be entrusted with any important task. The journeyman is qualified to perform almost any work done by members of the trade. An expert can perform quickly and with superior skill any work done by men in the trade. It is essential that the trade test differentiate between the journeyman and the apprentice, and the apprentice and the novice. Trade tests devised to make this classification are of three kinds: oral, picture, and performance. The oral tests are most generally used because they are of low

cost and may be applied to a large number of men in a comparatively short time without much equipment.

The following stages are characteristic of the usual method followed in obtaining a trade test.

1. Investigation and collection of trade data. An investigation is made into the conditions of the trade to determine the following facts. (a) The feasibility of a test in the trade, (in one instance it was found that the trade of gunsmith was not a recognized trade, though there were gun repairers). (b) The elements which require and permit of testing to determine whether men can be graded in the trade according to degrees of skill, (in some trades it has been found that the trade required simply the performance of a single set of operations and there were no gradations among the members of the trade). (c) The kinds of tests that can be used, (some trades such as truck driving are mainly matters of skill and for them performance tests are better than oral or written tests.) Other trades such as interior wiring and power plant operation are mainly matters of knowledge. For these trades oral and picture tests are better. After having discovered by inquiry that the trade is a recognized trade and can be tested, all the information necessary is collected in the field from all available sources, such as experts of the trade, trade union officials, literature of the trade, trade school authorities, employers and the like.

2. Compiling the questions. As a result of the information collected, a number of questions, usually forty to sixty are compiled, each of which calls for an answer which shows knowledge of the trade. The experience of the formulators of trade tests has shown that a good question meets the following requirements. (a) It must be in the language of the trade. (b) It must be a unit, complete in itself and requiring no explanation. (c) It is not a chance question that could be answered by a good guess. (d) It must be as short as possible and must be capable of being answered by a very short answer. (e) It must not be ambiguous.

3. Preliminary sampling. After the large number of questions originally formulated has been sifted down by the application of the requirements listed in the preceding paragraph the questions are used in a preliminary sampling on a number of tradesmen whose answers indicate the merit of the different questions and their grades of difficulty. In this



sampling tradesmen from different shops or plants are tried, in order to guard against specialized methods or modes of expression confined to a single locality.

4. Revision and formulation. The preliminary sampling affords a means of checking on the following points. (a) Is the test applicable to trade conditions? (b) Does the test represent good trade practice? (c) In what way can parts be profitably modified, supplemented or eliminated? (d) Does the test represent the whole range of the trade from the novice to the expert? (e) Is it a representative sampling of the whole range of trade processes? In the light of the answers to these questions the test is revised and then formulated.

5. Final sampling. Final sampling is made by testing a number of men, usually twenty, who are known to be typical members of each group (novice, apprentice, journeyman, expert.) Among the novices tested are usually some highly intelligent and mature men of good general knowledge but no trade ability. Statistical treatment of the results and of the answers to each question enables the determination to be made of a relative value of each individual question and the selection that makes a proper balance.

6. Evaluation. If a trade test is good, a known expert when tested, is able to answer all, or nearly all, the questions correctly; a journeyman is able to answer the majority; an apprentice a smaller part; and a novice practically none. This does not mean that each question should be answered correctly by all the experts, a majority of the journeymen, some apprentices and no novices. There are a few questions which show this result. Other types of questions, however, are more common. Some show a distinct line of cleavage between the novice and the apprentice. Novices fail, but apprentices, journeymen and experts alike answer correctly. There are likewise questions which are answered correctly by nearly all the journeymen and experts, but only a few apprentices and also questions that only an expert can answer. Each type of question has its value in a good test. The main requirement is that the tendency of the curve of distribution of the scores should be upward. A question that is answered correctly by more journeymen than experts or more apprentices than journeymen is undesirable and is at once discarded. A proper balance is made of the others.

7. Calibration. As each question is allowed a number of



points, usually four, it becomes necessary to determine how many points should indicate an expert, how many a journeyman, etc. This is accomplished by noting how many points were scored by the known experts and the known journeymen when they were tested. Ordinarily the expert scores higher than the journeyman and the journeyman higher than the apprentice. It frequently happens that a few journeymen score as high as the lowest of the experts and a few apprentices as high as the lowest of the journeymen. There are, consequently, certain overlappings between the classes. In calibrating, the object is to draw the dividing line between classes so that the overlapping shall be as small as possible. When these dividing lines, or critical scores have been established the test is ready to be applied.

Picture tests are made in practically the same way as the oral tests. The peculiar characteristic of picture tests is that the questions making up the tests relate to illustrations of trade tools and appliances.

Performance tests are fundamentally the same as the method of the vocational sample discussed earlier.

Certain important difficulties of the method of the trade tests may be mentioned. One is due to the fact that trade terms are very fluent, for instance the word "plug" probably means a score of different things in many trades, and probably no single one of these twenty "plugs" is universally known by that name. Localisms exist in various parts of the country such as "come along" in the linesman's trade in New York City. Foreigners, who compose no insignificant part of the total number of tradesmen, have a different nomenclature and one which is not usually a literal translation of the English. There is a difference in trade procedure in different parts of the country, carpenters trained in New York, Georgia and California will answer the question "What is done first in building a house" in very different ways and each in a way which is right according to his training. People may be coached, with ease, upon items of specific information—an individual with the barest trace of literacy can be taught to answer the following written questions: "What is your name?" "In what country were you born?" "In what state in the Union is your domicile?" Coaching is quite as easy in other specific information tests. Good trade tests determine simply the presence of trade experience and the degree of skill actu-

ally possessed. They are not prophetic tests which indicate probable aptitudes for trades which the person tested has never followed.

e. Miscellaneous empirical tests.

Finally, there are cases in which tests having vocational significance have been sought by purely haphazard and empirical ways. These may be designated as empirical vocational tests. This method consists of selecting at random a number of tests and obtaining measures of large groups of workers in any vocation and finding out if there are any tests which have a high positive correlation with the actual work. This is a strictly empirical method which is not based upon any *apriori* postulates as are most of the other methods. In the homely phrase of Carlyle, these tests seem to be "significant of much."

## 3. Procedure.

## A. Materials employed.

Nine tests were used, selected at random, from the group of tests which Woodworth and Wells (53) have standardized for the American Psychological Association. In groups 2 and 3 it would have been unwise to have attempted to use more than this number since it took approximately thirty minutes to test a subject. Under commercial conditions this was a great amount of time for the concern to give, and also for the subjects to give since they were losing their bonus for the time spent on the tests. The tests used are—opposites, verb-object, agent-action, action-agent, color naming, mixed relations, hard directions, number checking and form substitution.\* Of those tests in which more than one form is given by Woodworth and Wells the form which was used in this experiment is as follows.

\*Considerable time was spent by many subjects on the stimulus words "love" in the opposites test, on "kiss" in the verb-object test, and on "lungs" in the agent-action test. It was evident, that for some subjects, it was more difficult to respond to these words than to others of the list. If the tests were to be given over again these words should be taken out and others, of the same difficulty as the rest of the list, substituted for them.

Opposites	Verb-object	Agent-action	Action-agent	Mixed relations	
north	read	baby	gallops	Eye-see	Ear-
sour	tear	fire	bites	Monday-Tuesday	April-
out	throw	dog	boils	Do-did	See-
weak	paint	laborer	sleeps	Bird-sings	Dogs-
good	mail	pencil	floats	Hour-minute	Minute-
after	light	army	grows	Straw-hat	Leather-
above	sail	heart	sails	Cloud-rain	Sun-
sick	spin	pin	roars	Hammer-tool	Dictionary-
slow	lock	gun	scratches	Uncle-aunt	Brother-
large	wash	eyes	stings	Dog-puppy	Cat-
rich	bake	bird	shoots	Little-less	Much-
dark	spill	wind	melts	Wash-face	Sweep-
front	kiss	lungs	swims	House-room	Book-
love	polish	bell	explodes	Sky-blue	Grass-
tall	sweep	musician	aches	Swim-water	Fly-
open	fill	parrot	glows	Once-one	Twice-
summer	sharpen	clock	news	Cat-fur	Bird-
new	write	ax	cuts	Pan-tin	Table-
come	chew	broom	flies	Buy-sell	Come-
male	drive	mosquito	burns	Oyster-shell	Banana-

## B. Individuals serving as subjects.

## a. Group 1.

Seventy-seven young men and women, who were studying typewriting, stenography, and grammar in the Extension De-



partment of Columbia University, served as subjects in this first group. These people were divided among four different sections, three evening classes and one day class. In all of the evening classes it was impossible to obtain any adequate measure of the subjects' abilities, for they either dropped out in short time or failed to take the prescribed tests and examinations. The results worked up for the day class of forty-five students, forty-three female and two male are the only ones used in this investigation. The subjects made an intensive study of typewriting and stenography devoting two hours in class to each subject each day for five days a week. Some outside study was also done. The course lasted from the first part of October, 1915, until the middle of May, 1916. In group 1 the performances in the tests were correlated with abilities in typewriting, stenography and grammar.

b. Groups 2 and 3.

One hundred and eighteen typists, all female, in a large retail commercial concern in New York City were tested. Of these ten left the employ of the concern before any adequate measure of their abilities could be determined. The remainder were divided into two groups. Group 2 consisted of thirty-eight typists who had been working in the same division for at least ten months. Group 3 consisted of sixty-five typists who had been working in this same division for at least one month and a half and for not longer than six months. All of the subjects in both groups had been doing the same kind of typing. The records of five typists who had been working between six months and ten months were discarded. This division was made in order to allow, as well as possible, for the effect of practice in the particular kind of typing which the subjects in these two groups were performing, as it would have been unscientific to put all of the subjects in one group.\* In groups 2 and 3 the performances in the test were correlated with ability in typewriting only.

c. Age, education and experience of groups 1, 2 and 3.

The figures in the following table are in per cent of the number of individuals in each group.

AGE		Group 1.	Group 2.	Group 3
16-19 years old.		.46	.68	.91
20-24	"	.32	.32	.08
25-29	"	.06	.00	.01
30-34	"	.11	.00	.00
35-38	"	.04	.00	.00

\*See section 3, C, c, 2, (description of the work of groups 2 and 3).

EDUCATION				
Grammar school education		.02	.61	.47
1-3 years high school		.02	.36	.47
High School graduates		.51	.03	.05
1-3 years college		.22	.00	.00
College graduates		.22	.00	.00

EXPERIENCE				
Never worked previously		.85	.00	.31
Worked at typing 1-5 months		.00	.00	.10
Worked at typing 10-12 months		.00	.66	.26
Worked at typing more than 1 year		.00	.34	.23
Worked at other work than typing		.15	.23	.20

In the age table the mode for all groups is the group 16-19 years old. In the education table the mode for group 1 is the full high school education, for group 2 it is the grammar school education, and for group 3 it is either the grammar school education or the 1-3 years high school class.

### C. Methods of obtaining measures of abilities.

#### a. Stenography.

In February a mid-year examination was given in stenography which was graded by the A, B, C, D, F, method; A, B, C, being the order of satisfactory grades, D a condition and F failed. In each of these grades the mark was further qualified so that the A group, for instance, came to consist of A plus, A, and A minus. The individuals were then ranked in an order of merit in each of the three sub-divisions of each grade, for instance if there were three subjects in the A plus group they were ranked, according to their abilities, 1, 2, and 3. From these last rankings an order of merit series, from one to forty-five, was secured.

#### b. Grammar.

In February a mid-year examination was given in grammar, spelling, punctuation, paragraphing, letter writing, etc. This was a thorough three hour examination and the results were probably a good index to the subjects' abilities. The

grades in this examination were returned in per cent terms which were readily transformed into an order of merit series.

c. Typewriting.

More objective and reliable grades were obtained in typewriting than in either stenography and grammar, the subjects being measured in actual quantity and quality of output.

I. Group 1.

Each month a typewriting test was given by the instructor, in which the subjects copied a given piece of material on the typewriter, accomplishing as much as they could in ten minutes. The total number of words written in ten minutes, less five words for each error made, divided by ten gave the net number of words written per minute. This is an arbitrary method of scoring adopted as the best relationship between speed and accuracy by the International Typewriting Committee and has been incorporated into the international rules. The following table gives the number of subjects taking the examinations, the average net number of words written per minute, standard deviations, and ranges.

	No. of subjects	Av. no. words per min.	S. D. in words per min.	Range in words per min.
October	42	16.4	6.2	6. -37.
November	40	22.5	7.1	9. -51.
December	40	27.4	6.5	15.6-51.2
February	40	29.7	7.1	15.0-48.1
March	29	34.1	5.0	23.7-44.5
April	27	38.9	4.8	30.4-53.9

II. Description of the work of groups 2 and 3.

The commercial concern, in which the subjects of groups 2 and 3 worked, was composed, in part, of a number of stores which retailed merchandise and which received all of their orders by mail. The customers' orders, in the form received by the concern, were sent to the division in which these subjects worked, to have a sheet typed for each store in the concern from which the customer might have ordered a commodity. Thus an order would require anywhere from one to forty-six sheets written for it, as the customer might order articles, in one letter, from one to forty-six of the stores. A single order contained an order for a commodity from but one store and required but a single sheet to be written for it. Each sheet had typed on it the customer's name and address and specifications of the commodity ordered, such as



amount, size, shape, style, pattern, etc. This was a very particular kind of typing since there were practically no long phrases or sentences written and many figures, symbols and abbreviations were used. Any hierarchy of higher type-writing habits which the subjects might possess appeared to be of little advantage since it could not be made use of directly. All of the subjects in these two groups were working exclusively on the single orders.

It was ascertained, from the total number of words written and the total number of strokes made in several samples of three hundred sheets each that the amount of work done in a day could be expressed in terms of the number of sheets written in the course of a day's work. Thus the number of typed sheets per day furnished the measure of the amount of work turned out by the subjects. A group of "checkers" read each sheet typed by the subjects, compared them with the customer's original order and sent back to the subjects every sheet on which an error was made; and the subjects corrected the error in the course of the day's work, every error being thus penalized. A system was put into effect in which a girl circulated among the subjects and kept each subject supplied with orders, never permitting any subject to run out of orders. By this means no subject was able to select orders and thus receive easy ones.

A weekly bonus on the number of sheets written was in operation in which all of the subjects of group 2 shared almost every week and in which an average of approximately forty per cent of the subjects of group 3 shared. There was much evidence that each girl was urging herself to her best efforts.

### III. Group 2.

Records of the average number of sheets written per day were kept for ten months, from these there were selected the records of what appeared to be the best successive five weeks, the selection being based upon; 1. Attendance records; no one subject being absent more than once during this period, in which case she was credited, for the day absent, with the average number of daily sheets that she had written for the rest of the week in which the absence occurred; 2. Season of the year; the season being selected when the business was heaviest and when the subjects were being stimulated, by

their supervisor, to turn out as many sheets as possible. What appeared to be the best week's work was selected in like manner, the best week being a part of the five weeks which were selected. The best day's work was simply that day, within the five weeks' period, in which each individual subject did her best work. The five weeks' measures and the best week's measures were for the same calendar weeks for all subjects, but the best day's measures were not the same for all subjects. Thus three measures of ability were obtained: 1, daily average for five weeks; 2, daily average for one week; 3, best day's performance. The number of subjects, thirty-eight, was constant for each measure. These three cross sections of the subjects' abilities for different lengths of time were selected simply to ascertain whether the tests would correlate to the same degree with each. The following table gives the average number of sheets typed, standard deviations, and range of the average number of sheets typed.

	No. of sheets typed	S. D. of sheets typed	Range of sheets typed
Daily average for five weeks	450.3	64.8	360.2-662.0
Daily average for best week	509.5	68.0	402.5-735.2
Best day's performance	612.0	68.2	531. -812.

#### IV. Group 3.

Not all of the subjects of this group were working for the concern at the same time. Some subjects entered the employ of the concern two or three months after others and some of these latter had separated from the concern before the former had entered. The constant flux of the coming and going, of the arrival and departure of the individuals in this group, where the labor turnover was over three hundred per cent per year, made it impossible to take any one series of weeks' output, which weeks would be the same for all, as measures of abilities, as was done in group 2. A study of the records showed that many of the subjects began doing very good work about the fifth week after entering the concern. Consequently, the average number of sheets written per day during the fifth and sixth weeks of each individual's term of employment with the concern was taken as the measure of abilities. These weeks, then, were not the same calendar weeks for all individuals. The best day's work was also selected as a measure of abilities, the best day being that

particular day in which each subject did her best work. The following table gives the average number of sheets typed, standard deviations, and ranges. The number of subjects, sixty-five, was constant for both measures.

	No. of sheets typed	S.D.	Range
Daily average for 5th & 6th weeks	311.8	62.3	198.4-420.3
Best day's performance	361.9	65.3	258. -500.

#### D. Some remarks on the procedure.

The problem of securing groups in which the output of the workers is accurately comparable. In typing it is difficult to obtain a large group of typists in which all of the individuals of the group are performing the same kind of work which can be accurately measured, quantitatively and qualitatively. The difficulty lies in the fact that the group chosen at random is usually a heterogeneous one; that is, one individual does specific work which differs from the specific work of another individual, although both are performing the same generic act—typing. This difficulty has not entered into this investigation since all of the girls were a part of a homogeneous group and all were doing exactly the same kind of typing. The measures obtained in this investigation were not subject to the indeterminable error of a supervisor but were objectively obtained so that it was possible to state that one girl was so many points better or poorer than another girl.

Testing at various periods of the development of practice. In group 1 all of the subjects were tested when they had practically no practice or knowledge of typing. In group 2 all of the subjects were tested after they had acquired considerable practice and skill in typing. In group 3 the subjects were a mixed group being made up of individuals possessing varying degrees of practice and ability in typewriting.

Care taken to secure accurate measures of typing ability. More than one measure of typing ability was used in each group. In group 1 six measures, in group 2 three measures, and in group 3 two measures of ability were used. In group 2 approximately 9100 records of daily output, and in group 3 approximately 3000 records of daily output were tabulated and studied in order to obtain what appeared to be the best measures of typing ability.

Wide mental differences of the groups of subjects. The wide mental differences between group 1 and groups 2 and 3



in educational training indicate that two widely different classes in these respects were tested. It appears that the highest and lowest mental grades of girls in the typing profession were included in this experiment.

E. Administering the tests.

a. General procedure.

The subjects in group 1 were tested within two months after they had started their course. The subjects of groups 2 and 3 were tested after they had acquired considerable speed in typing. All of the tests in groups 2 and 3 and all of the tests but the number checking and the form substitution in group 1 were given individually. The number checking and the form substitution tests in group 1 were given as class tests. The method of conducting the experiment was simple. The subject was seated comfortably at a table, on which the blanks were presented one at a time, after it had been ascertained that the subjects thoroughly understood the instructions. In the number checking, form substitution and hard directions tests replies were filled in in pencil. In the remainder of the tests the replies were spoken. In the tests where the replies were spoken the experimenter, who sat next to the subject, had a copy of each test and if an inaccurate reply was made the subject was required to correct it before giving a reply to the next stimulus word, the subjects being told before hand that they would have to do this. In the number checking test each group was checked which contained any combination of the digits eight and nine.

b. Instructions to the subjects.

In the instructions to the subjects an effort was made to combine the proper comprehension of the experiment by the subject with an ideal uniformity of instruction. In the individual tests instruction was given by description, illustration, and execution. The subject was first clearly told the meaning of the test, then the experimenter performed a small sample of it, and finally the subject herself performed a small sample of the work. These samples were presented in typewritten form and the same samples used with all subjects. Six illustrations of each test were used with all subjects. In the class tests the subjects were simply told what they were to do and the experimenter performed a small sample of the work on the blackboard.

An attempt was made to impress upon the subjects the fact that speed was the main object of the tests. They were given a short talk on the applications of psychological tests and were told that their mental reactions were going to be timed and that their individual results were going to be compared with the results of the rest of the class. They were urged to go as fast as possible but were warned that they would either have to correct their mistakes as they went along or be penalized later for their mistakes (according to whether a class or individual test). It seemed that all the subjects tried to make all the speed possible.

c. Records.

In all the tests in groups 2 and 3 and in all the tests but the two class tests in group 1 the time was taken by a split second watch reading to fifths of a second. The watch was started when the experimenter judged that the subjects' eyes met the first figure or word on the test sheet, which was uncovered by the experimenter. In the class tests instructions were given that as soon as anyone had finished she was to raise her hand. When the first hand was raised the experimenter called upon the class to stop, having instructed them beforehand that he would do this. Thus the time for all was the same and the results were worked up for the amount accomplished.

4. *Results and how obtained.*

## A. Scores in the tests.

The following table gives the scores made in the tests for groups 1, 2 and 3.

	Group 1.		Group 2.				Group 3.			
	Av.	S.D.	Av.	S.D.	X	Z	Av.	S.D.	X	Z
Opposites	28.56	8.26	39.00	12.25	.36	.48	43.37	16.12	.52	.95
Verb-object	31.43	7.91	38.59	10.26	.23	.29	37.23	10.39	.18	.31
Agent-action	31.92	10.09	40.61	16.21	.27	.60	39.81	14.28	.25	.41
Action-agent	33.06	6.75	46.33	11.83	.39	.75	50.19	17.35	.51	1.57
Mixed relations	76.03	28.65	99.47	34.30	.31	.20	101.00	34.71	.33	.21
Color naming	53.81	8.99	62.09	8.17	.15	+.09	62.07	9.80	.15	.09
Hard directions	110.01	28.21	186.25	60.00	.69	1.13	188.72	53.91	.71	.91
No. checking	151.52	28.27	157.03	34.72	.04	.23	176.03	33.52	.16	.19
Form subs.	124.40	13.11	150.92	21.70	.21	.65	152.58	30.46	.23	1.32
No. c. Errors	4.70	3.38	8.20	7.62	.74	1.25	7.95	7.83	.68	1.32
Form s. "	.48	.67	2.63	4.48	4.48	3.85	3.17	3.22	5.46	3.81
Hard d. "	1.64	.45	4.33	2.48	1.69	4.51	4.00	2.96	1.44	5.58
No. of subjects	45		38				65			

Av.—average score of the group in seconds and hundredths of a second, in the tests from opposites down through form substitution.\* In the last three tests not the time but the average number of errors and hundredths of an error is given.

X—percentage that the average of the group is above or below the average of group 1. A + quantity indicates that the average is below the average of group 1, or a better performance, and no sign before the figures indicates that the average is above the average of group 1, or a poorer performance.

Z—percentage that the standard deviation of the group is above or below the standard deviation of group 1.

\*Since in group 1 the grades in the number checking and form substitution tests are measures in terms of the amount accomplished (average in number checking—117.31, average in form substitution—76.31) the figures, in the table under group 1, for these two tests, are roughly derived, for means of comparison in this table only, by dividing the time (constant for the group in each case) by the amount done and multiplying by the number of groups of figures in the number checking test, and the number of figures in the form substitution test.

$\frac{\text{time}}{\text{amt. done}}$  no. of cases.

Groups 2 and 3 showed themselves to be greatly inferior to group 1 in performing the tests. Groups 2 and 3 were made up of girls who were working long hours for very low wages



and who came from families in poor circumstances. Twenty-four per cent of the subjects of group 2 and twelve per cent of the subjects of group 3 came from homes where a foreign language was spoken. The average foreign-born performance, in each test, in each group, was compared with the average performance of the rest of its group and was found to be higher by no more than six per cent in any one test. The inferiority in performing the tests in comparison with group 1 could, therefore, not be traced to difficulty in using the English language due to speaking a foreign language in the home.

#### B. Treatment of errors.

At the outset of the experiment it was planned to treat separately the time required to perform the test and the number of errors made and to correlate the number of errors made in the number checking, form substitution, and hard directions tests with the grades in stenography, grammar, and typewriting. The standard deviations of the number of errors made in each of the tests being in every case almost as high if not higher, than the average number of errors made, no attempt to correlate the errors with the work has been made since the number of errors made is so unreliable as a measure.

#### C. Methods of combining the measures.

When several tests have been made of an individual's abilities it is often desirable to show the success of the individual in the series of tests taken as a whole. A good way of doing this is by the order of merit method. The measures of the individuals are so arranged that it can be stated that an individual stood thirteenth from the top in one test, twenty-third in another test, eighteenth in still another test and that her average rank was eighteenth. This method was used in correlating performances in the tests with ability in stenography and grammar.

Although the order of merit method is rather a rough method it has worked well with certain kinds of material in the past. But to transmute a series of quantitative measures into an order of merit series is to throw away a great deal of information contained in the series. Another method consists in taking the average of the group as zero and expressing the individual's standing as a devia-

tion above or below this average. The measure of deviation, usually the standard deviation, is taken as the unit and all deviations are expressed as multiples of this unit. What this method does is to assign to each individual a quantitative position in the distribution of the group. This method was used in getting reduced measures in the tests which were correlated with ability in typewriting in groups 2 and 3 and with grades in typewriting in group 1.

#### D. Correlations.

##### a. Stenography and grammar.

#### I. Method used.

The rank differences formula was used in correlating the tests with abilities in stenography and grammar,

$$r=1-\frac{6\sum d^2}{n(n^2-1)}$$

#### II. Correlations.

The following table gives the correlations with the tests and abilities in stenography and grammar.

	Steno.	Grammar
Verb-object .....	.36	.37
Number checking .....	.07	.22
Color naming .....	.34	.38
Action-agent .....	.23	.35
Agent-action .....	.19	.37
Form substitution .....	.40	.16
Hard directions .....	.46	.54
Mixed relations .....	.31	.43
Opposites .....	.45	.40
No. of subjects .....	45	45

##### b. Typewriting.

#### I. Method used.

In correlating the tests with abilities in typewriting, in the three groups, a derivation of the Pearson formula was used (52),  $r=2Av\left(\frac{a+b}{2}\right)^2-1$  a and b being the reduced measures in the test and typewriting. The following table gives the correlations of the tests with abilities in typewriting.

## II. Correlations.

	Group 1.							Group 2.			Group 3.	
	October	November	December	February	March	April	Average Oct. Nov., Dec., Feb.	Five weeks	Best week	Best day	5th & 6th wks.	Best day
Verb-object	.41	.43	.46	.55	.21	.57	.46	.28	.32	.34	.33	.35
No. check.	.45	.47	.37	.53	-.01	.30	.45	.28	.34	.41	.34	.39
Color naming	.30	.43	.55	.41	.29	.61	.40	.39	.39	.40	.36	.38
Action-agent	.42	.43	.29	.41	.00	.40	.39	.13	.29	.26	.29	.28
Agent-action	.29	.19	.40	.31	-.02	.28	.30	.02	.20	.24	.18	.26
Form subs.	.21	.27	.11	.37	.30	.42	.24	.04	.14	.17	.17	.20
Hard dir.	.11	.14	.19	.13	.34	.32	.14	-.20	-.20	-.08	.16	.19
Mixed rel.	-.09	.21	.17	.25	.04	.00	.13	-.07	-.12	-.09	.03	-.04
Opposites	.17	.11	.07	.15	.19	.54	.12	-.09	-.09	-.11	-.07	.06
No. of subj.	42	40	40	40	29	27	40	38	38	38	65	65

## E. Discussion of correlations.

a. Significant tests for typewriting and significant tests for stenography and grammar.

In comparing the correlations obtained between tests and typewriting with the correlations between tests and stenography and grammar, if a line is drawn horizontally across the paper in the middle of the list of tests, in each of the preceding tables, it will be observed that, in most cases, the tests which correlated best with stenography and grammar and worst with typewriting lie below this line and the tests which correlate best with typewriting and worst with stenography and grammar lie above this line. The two tests which are not language tests, number checking and form substitution, do not correlate well, except in one instance, with stenography and grammar, but one correlates very well and the other fairly well with typewriting.

## b. Typewriting, group 1.

In considering the first four months' correlations in typewriting one is met with the uniformity of the results for any one test during those four months. If a test correlates well with one month's measures of abilities in typewriting, it can be predicted that it will correlate well with another month's measures; if a test correlates poorly with one month's measures it can be predicted that it will correlate poorly with another month's measures.



The March correlations with the verb-object, number checking, agent-action and hard directions tests show no correspondence with the correlations for these tests for the first four months. This can be accounted for as follows: The March typewriting test was an especially difficult one which probably means that it called for reactions which the subjects had not formed the habits of making or that it called for the functioning of processes which did not exist or which were not well developed. Since the exercise was exceptional it could be expected that the results of the exercise would be exceptional and it is quite possible that the correlations would be exceptional and that the uniformity of correlation would be absent. It will also be observed that the number of subjects dropped from forty to twenty-nine this month.

The April correlations show a good degree of correspondence with the correlations for the first four months except for the form substitution, hard directions, and opposites tests. Here again it will be observed that only twenty-seven subjects took the exercise. This reduction in the number of subjects taking the exercise does not mean entirely a process of the elimination of the poorer typists; some of the poorer typists had dropped out of the course and some of the best typists were in the University offices working as typists. Inspection of the table will show that in the first three tests in the list, those which correlate highest with typewriting, and correlations are usually higher each month as the subjects acquire more speed in typing.

c. Typewriting, groups 2 and 3.

In the tests which correlate well in group 2 the correlations with the best week's measure of ability are higher than the correlations with the five weeks' measure of ability, and the correlations with the best day's measure of ability are higher than either the correlations of the five weeks' or the best week's measure of ability. In the same tests in group 3 the correlations with the best day's measure of ability are higher than correlations with the fifth and sixth weeks' measure of ability. The best day measures the ability when a spurt is made and the best spurt over a long period of time is probably the best index of a subject's ultimate ability that can be found in practice of groups of this sort. Since the correlations are higher in these short intense periods it is concluded

that the tests tend to measure the physiological ability in typewriting.

d. Typewriting, groups 1, 2 and 3.

The seventh column in the table, the average of the correlations for October, November, December and February was obtained as a means of comparison of group 1 with groups 2 and 3 rather than comparing any one month's correlations of group 1 with the correlations of group 2 and 3. In groups 2 and 3 the correlations, for the first four tests in the list, in every case, are lower than those of group 1. With slight exceptions the tests which are correlated well in one group are correlated well in other groups and the tests which are correlated poorly in one group are correlated poorly in other groups.

#### F. Partial correlations.

The four tests which were correlated highest with typewriting ability were used in deriving a multiple correlation with the measures obtained in the typewriting exercise of February in group 1.

The calculation of the regression equation with five variables is,—Measure in typewriting = .3721 (measure in verb-object) + .0109 (measure in color naming) + .0848 (measure in number checking) + .0852 (measure in action-agent). Or for practical purposes, after multiplying by 24.3, a convenient constant; Measure in typewriting = 9 (measure in verb-object) + .3 (measure in color naming) + 2. (measure in number checking) + 2 (measure in action-agent). After performing the arithmetic this gives a correlation of .61 with a probable error of .07.

5. *Applicability of the results.*

The correlation method has been used to measure the degree of association between performance in the tests and typing ability and it is now necessary to obtain the line of regression or the line of best fit in order to predict, for any individual in the group, her most probable performance in typing from her performance in the tests.

The accompanying diagrams, appended hereto, show frequency distributions in groups 1, 2 and 3 of the reduced measures in the combination of the four tests, as obtained by the weights assigned by the method of partial correlation, and the reduced measures in typewriting, the former being plotted on the x axis and the latter on the y axis. From the general sweep of the scatter diagrams, it is evident that, as the measures in typewriting increase or decrease the measures in the combination of the four tests increase or decrease. The straight line AB, giving the best fit to the data, is fitted to the points in the diagrams by means of the formula  $y=mx+b$  which for group 1 becomes  $y = .68x + .00b$ .\*

The expression best fit is seldom defined. Its significance varies with the problem in hand and it generally means a fit which is convenient and which for the problem to be solved, gives satisfactory results. The principle upon which the values of m and b are determined is so to choose m and b as to make the mean square deviation of the observations from the resulting straight line a minimum. The pertinancy of this principle for the problem of prediction is plain since when observations are distributed according to the normal law, the Tables of the Probability Integral enable us to com-

\* $y = mx + b$  or  $y - \bar{y} = r \frac{\sigma_y}{\sigma_x} x - \bar{x}$  where  $\bar{y}$ ,  $\bar{x}$  and  $\sigma_y$ ,  $\sigma_x$  are the means and the standard deviations, respectively of the y's and x's.

$$y - .00 = .61 \frac{109.5}{96} x - .00b.$$

$$y = .68x + .00b.$$

pute the probability of a deviation equal to any multiple or submultiple of the root mean square deviation. Moreover, as in all problems of prediction it is desirable to have the root mean square deviation as small as possible it is obvious that a straight line which fits given data so as to make the mean square deviation of the points from a straight line a minimum is, for the problem of predicting one variable from a knowledge of the other, a good fit to the data.



We have found the best fitting straight line connecting performances in the tests and ability in typewriting by means of the formula  $y=mx+b$ . For any given value of  $x$ , representing a performance in the tests, we can predict, by means of this formula, the most probable performance which will be achieved in typing. The problem is not completely solved by finding the equation connecting the two variables. We are not content to predict the most probable values of  $y$ , but we wish to know, in addition, the degree of accuracy of the predictions. The scatter formula supplies an answer to this latter question. The formula  $S$  (scatter)  $= \sigma y \sqrt{1-r^2}$  (30)\* shows the *accuracy* with which, using the equation to the best fitting straight line, the magnitude of the one variable may be predicted from a knowledge of the other. Substituting the values, for group 1, in the formula gives  $S=84$ .† When the association between the two variables is perfect, or when  $r = \pm 1$ ,  $S = 0$  and all of the points lie on the straight line, and by means of the equation to the best fitting straight line, we can predict exactly the value of  $y$  corresponding to a given value of  $x$ . But it is an exceedingly rare case when  $r = \pm 1$ , and when  $r$  lies between these two limiting values, we can still predict the results with a knowledge of the probabilities in favor of the prediction. When the association between the two variables is not perfect,  $r$  falls between the limiting values  $\pm 1$ , and the equation to the best fitting straight line enables the computation of the most probable value of  $y$  corresponding to a given value of  $x$  and the value  $S$  shows the accuracy with which, using this equation to the best fitting straight line, the magnitude of the one variable may be predicted from a knowledge of the other or enables us to say within what limits any proportion of the measures are scattered about the straight line. The degree of association between performance in the tests and typing ability is measured by the correlation formula whose value is  $r = .61$ : the formula by means of which typing ability can be predicted from performance in the tests is  $y=.68x+.00b$ : and the error of the predictions is measured by means of the scatter formula the value of which in this case is  $S = 84$ .

It is known from the Table of the Probability Integral that when the distribution of the points about the straight line

\*For the derivation of this formula the reader is referred to Chapter 2.

† $S = \sigma y \sqrt{1-r^2} = 106.5 \sqrt{1-.61^2} = 84$ .

is normal, 99.7 per cent of the observations lie within a deviation from the straight line equal to  $\pm 3S$ , 95 per cent lie between  $\pm 2S$ , and 68 per cent lie between  $\pm S$ . Since  $S = 84$ ,  $2S = 168$ . In the chart for group 1 the lines  $\pm 2S$  are drawn parallel to the line AB and 168 points above and below it, respectively, on the  $y$  axis.

The accuracy of the forecasts is measured by  $S = \sigma_y \sqrt{1-r^2}$  where  $\sigma_y$  is the standard deviation of a concrete series, and  $r$  is the correlation between the two series.  $S$  measures the accuracy of the predictions because it shows how the prediction formula enables us to reduce their variability. If there were no predicting formula the variability of the series that we wish to know would be  $\sigma_y$  but by the use of the formula the variability of the predictions is only  $\sigma_y \sqrt{1-r^2}$ . The factor  $\sqrt{1-r^2}$  measures the reduction in variability that is gained by means of the predicting formula. If, therefore, we wish to compare the accuracy of predictions of two different series,  $\sigma_y$  being constant, the measure of the relative accuracy is given by  $\sqrt{1-r^2}$  and the smaller the value of  $\sqrt{1-r^2}$ , the greater the accuracy of the predictions. The same idea may be put in a different way by saying that the greater the value of  $r$ , the greater the accuracy of the predictions.

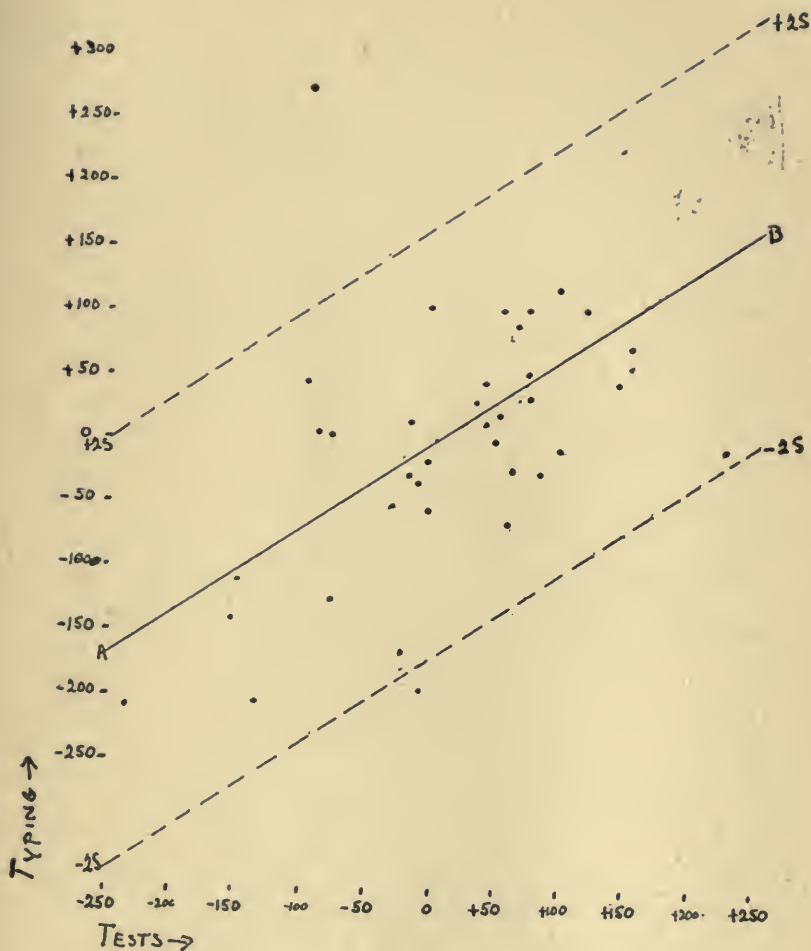


Fig. 1

Scatter Diagram for Group 1.  $r = .61$ .  $y = .676x + .00b$ .  
 $2s = 168$ . Pop. = 40.



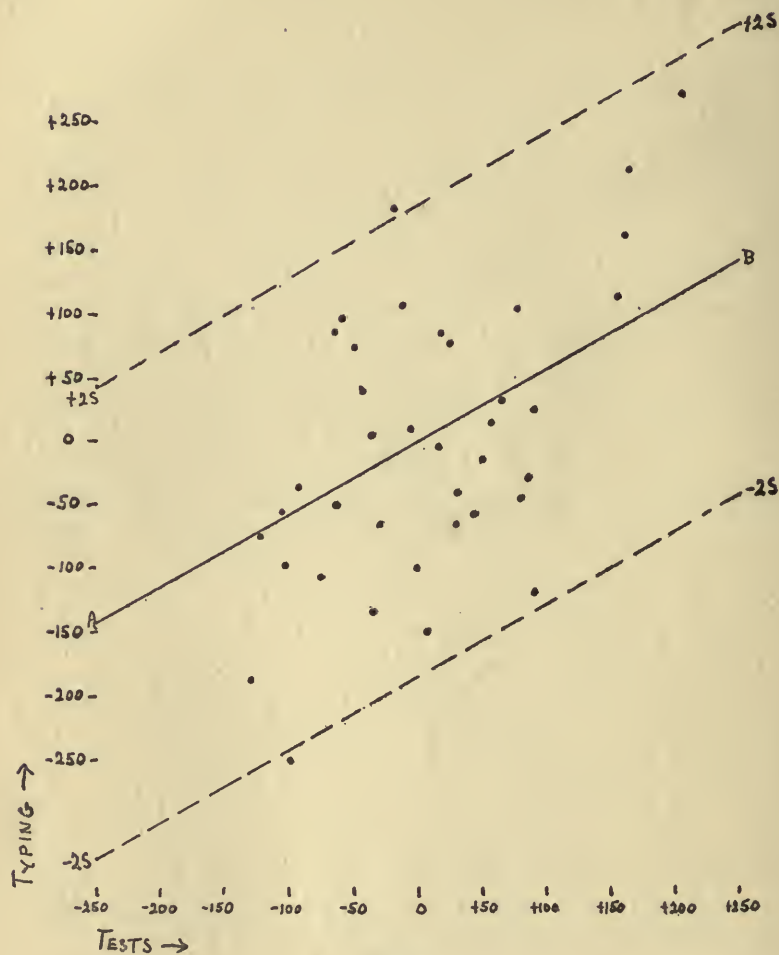


Fig. 2

Scatter Diagram for Group 2.  $r = .47$ .  $y = .59x + .00b$ .  
 $2s = 186$ . Pop. = 38.

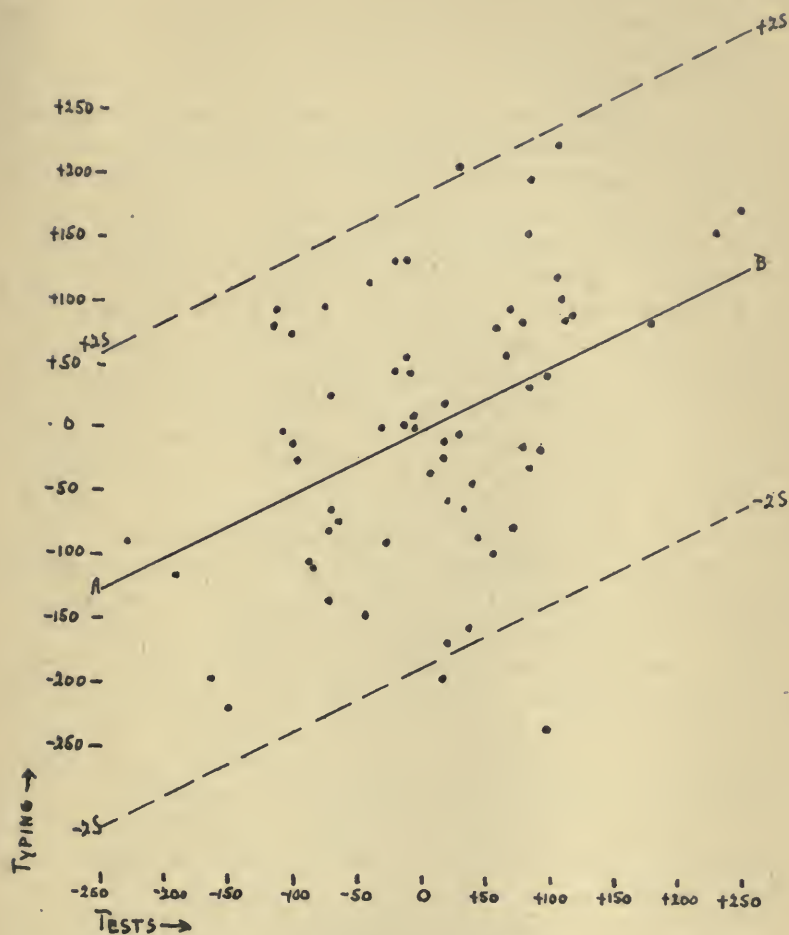


Fig. 3

Scatter Diagram for Group 3.  $r = .45$ .  $y = .515x + .00b$ .  
 $2s = 187$ . Pop. = 65.

In group 1 all but five subjects received passing grades in typewriting which, in terms of the diagram, means that all subjects scoring above  $-150$  on the  $y$  axis accomplished satisfactory work. We will take this point, for the group, as being the lowest point to which any one subject may reach and still do satisfactory work in typing, then it can be predicted with 95 per cent accuracy that all the subjects who score  $+25$  in the combination of the four tests will score between  $-150$  and  $+185$  in typewriting or will accomplish satisfactory work in typewriting (intersection of the line  $-2S$  with  $+25$  on the  $x$  axis  $= -150$  on the  $y$  axis and the intersection of the line  $+2S$  with  $+25$  on the  $x$  axis  $= +185$  on the  $y$  axis). Or for any specific individual such as one, for instance, who receives a score of  $+225$  in the tests, we can predict that she will be able to accomplish a score of  $+150$  in typewriting (intersection of the line of best fit with  $+225$  on the  $x$  axis  $= +150$  on the  $y$  axis) and we know with 95 per cent accuracy that her accomplishment in typewriting will not be lower than  $-12$  nor higher than  $+304$  (intersection of the line  $-2S$  with  $+225$  on the  $x$  axis  $= -12$  on the  $y$  axis and the intersection of the line  $+2S$  with  $+225$  on the  $y$  axis  $= +304$  on the  $y$  axis). For convenience the reduced measures have been used. After the arithmetic of balancing the scores in the tests has been computed the same thing may be accomplished with the actual scores in the tests and in typewriting.

There appears to be a great deal of difference in the opinions of psychologists as to just how much of a coefficient of correlation, with its probable error, constitutes a significant correlation, some psychologists having reported figures lower than  $.25$  as being significant coefficients of correlation. Some of this confusion can be cleared up by means of this method, and the practical significance of coefficients of correlation may be determined.

The problem of vocational guidance is to give an individual counsel as to what sort of an occupation or profession his potential capacities fit him for, so that he can be definitely informed that he will be successful in certain lines of endeavor and unsuccessful in certain other lines, or in more specific cases for any one occupation such as assembling, typewriting, teaching psychology, etc., whether the individual will be a success or a failure and to just what degree he will be a success or a failure in that specific occupation. Unless the chances



of an individual succeeding or not succeeding are very accurately determined it does the individual very little good to know that his chances for success are such and such a percent, as seventy-nine or eighty-two percent, etc. For the problem in hand, in the typewriting occupation, this sort of counsel given to an individual may result in her spending an amount of time ranging from a few months upward in training and then she may finally learn that she is one of those unfortunate ones outside of the seventy-nine percent and that she will never become a successful typist. It is evident that of ten individuals examined if we can predict that the chances for success in each case is, say, approximately eighty-nine percent and they are all counselled to take up the study of typewriting and go through a period of training and at the end of the training period nine develop into successful typists and one becomes a failure at typing then as far as this one individual is concerned she has been misguided and her time and expense of training has been wasted to say nothing of the effect that the failure may have on her temperament and character.

Vocational guidance should mean accurate scientific prediction for each individual and if, in the future, applied psychology, in the field of vocational guidance, would escape the many abuses that have been heaped upon it in the past it will not attempt to establish a system of vocational guidance until it can be almost perfectly sure of each and every one of its predictions. To establish such a system of vocational guidance which is to make no grave errors it will probably be necessary to have a coefficient of correlation well above ninety with a very small probable error and with a small scatter value; just what coefficient of correlation, probable error, and scatter value will be necessary will probably be impossible to accurately determine until the actual work has been accomplished and the scatter diagrams drawn and interpreted.

From the results of this research, it is interesting to know that we can predict that individual A who scores 140 in the psychological tests will most probably attain an ultimate ability in typing of forty words per minute, and in 99.7 percent of the cases she will attain an ultimate ability in typing varying between not more than thirty-one to forty-nine words per minute, and that individual B, who scores ninety-six in the psychological tests will probably attain an ultimate ability in

typing of thirty-two words per minute and in 99.7 percent of all such cases she will attain an ultimate ability in typing varying not more than between twenty-three to forty-one words per minute. But it will readily be seen that these predictions are of very little practical value for the individual in that they afford her very little definite information about just what ultimate ability she, individually, will attain. From our data we can readily predict that in all probability A will become a better typist than B but still it is quite possible for A to develop into a poorer typist than B. The results of this research show, then, that the correlations are not high enough and the scatter value is not small enough to warrant the establishment of a system of vocational guidance whereby an individual can be advised, with any sufficient degree of safety, to undertake or not to undertake to prepare herself to become a typist.

The problem of vocational guidance, then, is supplied by the individual who wants to know of what his potential capacities consist and what sorts of occupations he is fitted to follow successfully and what sorts of occupations he is not fitted to follow successfully. The problem of vocational selection is of an other sort, here the need is supplied by the profession or the industrial concern and not directly by the individual. The problem of vocational selection is—given a certain number of applicants in excess of the number required for a specific job, to select those who, *on the whole*, are best fitted to succeed. In practical life the employment manager has a number of vacancies open for a specific job and of a limited number of applicants for that job his problem is to select from the applicants those who, on the whole, are most likely to succeed. If ten jobs in the typewriting department are open and twenty typists are induced to apply for the jobs the problem is to select the ten who, on the whole, will do the best work. If this can be accomplished there will be an ideal state of affairs. However, this ideal has never been reached in any system of vocational selection. If we have something which works with some degree of success, with such a degree of success never before attained by any other method, which has been checked up scientifically, then we are justified from a pragmatic view point in using it until a better method appears. If we can have a system of vocational selection which can pick from the twenty applicants (ten poor ones and ten good ones)



a majority of the good ones and weed out a majority of the poor ones we have something which works and which is valuable. If by our method of selection of ten typists from twenty applicants eight finally turn out to be very good typists, one a typist of medium ability, and one a poor typist then we have erred to some extent but on the whole we have made a selection which is a good one. By pure chance we would probably have selected five good typists and five poor typists and practically all of the traditional methods of selecting typists appear to do nothing more than this even if they accomplish this much.

It is evident from the foregoing data that this method of empirical vocational tests gives a far more reliable criterion of vocational selection in the field of typewriting than has been obtained by any other method. If a commercial concern finds it necessary to employ a number of typists a test in typing may be given, as so many concerns do, which may or may not determine the actual ability in typing at the moment. This will insure the concern that they are not hiring people who, at the moment are wholly inefficient. However, if we are looking forward to the ultimate ability which the typists will achieve it is not safe to hire a person to do typing on the basis of the performance of a sample of typing for the typing ability of a person when she first enters the employ of a concern is often no index of the ability that she may develop.<sup>†</sup> Along with the test in typing the team of tests may be administered and from the results the employer may predict with 95 percent accuracy or with 99.7 percent accuracy, or with any other degree of accuracy that is required whether or not the applicants will develop into good typists and in addition within just what limits their capacities in typing will enable them to develop in this field, these predictions being made in definite quantitative terms such as the net number of words written per minute.\* The prediction of ability by means of the tests will do injustice to some individuals but it will vindicate it-

---

<sup>†</sup>In group 3 there was a coefficient of correlation of .28 with a probable error of .10 between the first day's output and the best day's output, the latter measure being taken within a period of six weeks.

\*The supply of typists in New York City in the last few years has been sufficient to enable a high degree of selection to be made. During the time in which the experiment was carried on in the commercial concern, in 1917, applications were received from two hundred and eighty-two typists for ninety-six positions which were open during that period.



self by the percentage of best applicants, and the percentage of poorest applicants which it will be able to detect.

Some concerns require an average of twenty-two words of connected discourse to be written per minute, others require thirty-five words, others forty-three words, and so on. Some concerns require twenty words of unconnected discourse to be written per minute, other concerns require ninety-five strokes of indexing work to be written per minute and so on. There are then various degrees of quantitative and qualitative typing required by various concerns and a typist who is capacitated to satisfy the requirements of one concern may not be capacitated to satisfy the requirements of another concern. The tests have developed a definite scale of efficiency in typing, from lower to higher grades. Since each of the various concerns requires a different average ability in its typists, higher or lower than that required by other concerns, it will be necessary, before the tests can be adequately applied, first to determine this required standard for the given concern. Reference can then be made to the scale and a standard time for performance in the tests can be fixed upon. Candidates may then be selected in order, by their grades, from the highest scores down, until the required number have been secured. The employment manager may stop at the "minimum score" in the tests corresponding to the minimum standard of efficiency required for the job and then attempt to procure a new group of applicants in order to secure applicants who can pass the tests with better than minimum grades, or if the supply of possible applicants is limited, he will at least have clear information, at the start, as to what he may expect of those whom he is forced to select from the group who obtain inadequate scores in the tests. Even here he is able to make the best of a bad situation.

## APPENDIX

## 1. The supervisor as a judge of abilities.

In correlating abilities in performing tests with other abilities it is often impossible to obtain objective measures, of individuals in a group, in exact terms of output or of a definite quality and quantity of a particular capacity or ability, in which case measures of abilities are estimated by a teacher or supervisor. The following analysis of a particular supervisor's estimate of abilities may contribute, in some small way, to the reliability which is to be attached to such measures.

The supervisor of the work of groups 2 and 3 was asked to arrange the subjects in each group in an order of merit series for typewriting ability. Emphasis was placed upon the necessity of the supervisor basing his judgments on typewriting ability only and to exclude all other factors. The supervisor was in possession of a copy of all the records of the output. The following table gives the correlations of his judgment with the measures of abilities according to output.

Group 2.	r.
Daily average for five weeks	.74
Daily average for best week	.61
Best day's performance	.59
Group 3.	
Daily average for 5th & 6th weeks	.53
Best day's performance	.42

The correlations tend to show that the supervisor was a better judge of ability which manifested itself over a long rather than a short period of time. He was either a poor judge of ability which manifested itself in spurts rather than ability which manifested itself in steadiness from day to day, or he based his estimates on something else than pure ability, or both alternatives may be true. In conversation with the supervisor, he realized that there were other things which he had taken into consideration in making the judgments, although he tried to judge for actual typing ability only, such as steadiness and general reliability of output, attendance, conduct, ability of the subject to make herself an agreeable member of the group, etc. In other words he was incapable of judging actual typing ability and his estimates were based on the general desirability of the subject to the concern.

This is a problem which every investigator who obtains measures of abilities by means of some one's estimate of those abilities, must confront. It is probably impossible for the general run of supervisors to grade for pure ability and to omit all other factors. This may be a favorable point if tests are sought which will be indices of an individual's total value on a job.

2. Averages of the scores in the tests arranged according to age and education.

In the following tables group M is made up of all the subjects who were tested in the Extension Department of Columbia University and group N is made up of all the subjects who were tested in the industrial concern. The number of subjects in each sub-division is small and the deviations from the average are high, so that, a good amount of caution must be exercised in any interpretation.

When all the evidence is evaluated no difference with age and performance in the tests is found.

Since the comparison, in group M, in education, is made between college graduates, high school graduates, and grammar school graduates, only those subjects who had reached their twenty-second year are used in the group. The tables show that there is a difference between education and performance in some of the tests, but where there is a difference it does not appear to be strongly marked.



## AGE

## Group M

Age	Subjects	Opposites	Verb-object	Agent-action	Action-agent	Color naming	Hard directions
16	1	35.8	30.6	29.4	44.6	51.0	107.4
17	3	36.33	32.60	42.20	37.07	55.73	109.47
18	14	28.47	31.38	33.54	34.70	55.04	118.10
19	12	31.42	34.28	35.53	35.35	54.03	117.67
20	8	23.68	30.65	29.30	32.45	52.65	110.75
21	8	29.78	28.50	28.65	31.50	56.15	98.45
22-25	10	25.56	28.06	26.74	30.42	50.14	89.86
26-30	9	25.69	33.53	30.18	34.42	56.18	112.04
31-44	12	28.08	36.19	29.92	37.12	52.52	132.98

## Group N

16	13	39.78	39.32	43.86	51.52	62.20	176.20
17	40	38.92	36.79	38.35	47.16	62.21	177.40
18	28	43.00	37.62	39.23	51.07	58.13	196.92
19	14	41.13	35.74	35.47	39.79	63.93	195.87
20	5	42.88	37.36	33.72	47.24	65.04	183.72
21	15	44.40	38.88	41.51	48.36	62.33	176.60

## Groups M &amp; N combined

16	14	39.50	38.70	42.83	51.03	61.40	171.28
17	43	38.74	36.49	38.62	46.46	61.67	172.66
18	42	38.14	35.54	37.33	45.61	57.10	170.65
19	26	36.65	35.07	35.51	37.77	59.36	159.75
20	13	31.06	33.23	31.00	38.14	57.41	138.82
21	23	39.31	35.27	37.03	42.49	60.19	149.42

## EDUCATION

## Group N

Education							
Gram. sch. grad.	54	44.51	39.12	41.74	51.03	63.12	186.76
½ yr. high sch.	9	37.40	32.88	35.07	43.87	62.09	190.31
1 yr. high sch.	21	35.81	39.01	35.35	41.43	59.17	187.62
2 yrs. high sch.	19	42.15	36.31	43.25	48.82	62.51	195.14
3 yrs. high sch.	8	35.50	32.05	37.05	50.38	59.82	154.53
4 yrs. high sch.	4	37.70	37.75	37.45	50.20	64.15	158.60
combining,							
Gram. sch grad.	54	44.51	39.12	41.74	51.03	63.12	186.76
½-1 yr. high sch.	30	36.29	37.17	35.27	42.16	60.05	188.43
2-4 yrs. high sch.	31	39.28	35.39	40.90	49.40	62.01	179.94

## Group M

Gram. sch. grad.	8	34.04	30.64	38.84	45.36	59.88	180.08
High sch. grad.	14	26.23	31.30	27.54	33.41	50.93	107.64
College grad.	11	25.11	26.47	25.82	30.87	52.44	95.93

3. Sample of an Army Trade Test. (12).\*

Welder, Cutter  
Cutter, Oxy-acetylene Operator

Oral

1. Q. What is it called when the edges of metal sheets are welded together in different spots to hold them in place for welding?  
A. Tack (spot). Score 4.
  2. Q. What chemical is mixed with water to form acetylene gas?  
A. Carbide. Score 4.
  3. Q. What metal do you use to braze brass and cast iron together?  
A. Brass (bronze). Score 4.
  4. Q. What will happen if oil gets on the oxygen regulator or hose?  
A. (1) Explode (blow up). Score 4.  
(2) Catch fire (burn up). Score 4.
- \* \* \* \*
18. Q. What is put in the acetylene tank to prevent explosions?  
A. Acetone. Score 4.

Rating the candidate

19 and below .....	N
20 and 21 .....	A+
22 to 35, inclusive .....	A
36 and 37 .....	A+
38 and 39 .....	J—
40 to 55, inclusive .....	J
56 and 57 .....	J+
58 and above .....	E
There is no E— or E+ rating.	

\*Vol. 2, p. 148-150.

## BIBLIOGRAPHY

1. American Academy of Political Science, Annals of, Personnel and Employment Problems. 1916. Phila. Stabilizing Industrial Employment. 1917. Phila. Steadying Employment. 1916. Phila.
2. Baerwald, Richard. Zur Psychologie der Vorstellungstypen. Schriften der Gesellschaft für psychologische Forschung. Leipzig, J. A. Barth, 1916.
3. Boas, Franz. Determination of the Coefficient of Correlation. Science, N. S., 29, 823-824.
4. Boas, Franz. The Measurement of Variable Quantities. Science Press, N. Y., 1906.
5. Book, William F. The Psychology of Skill. University of Montana Studies in Psychology, I, 1908.
6. Brown, H. C. Language and the Association Reflex. Jour. of Phil. Psy. & Sci. Meth., 13, 1916, 645-649.
7. Brown, William. The Essentials of Mental Measurement. Cambridge University Press, 1911.
8. Bryan, W. L. & Harter, N. Studies in the Physiology and Psychology of the Telegraphic Language. Psy. Rev., 4, 1897, 27-53 & 6, 1899, 346-375.
9. Burt, Harold E. Professor Munsterberg's Vocational Tests. Jour. of Appl. Psy., I, 1917, 201-213.
10. Chapman, J. Crosby. The Learning Curve in Typewriting. Jour. of Appl. Psy., 3, 1919, 252-268.
11. Chapman, J. Crosby. Army Trade Tests. Henry Holt & Co., 1921.
12. Chapman, J. Crosby & Hills, Myra E. Positive Acceleration in the Improvement of a Complex Function. Jour. of Exper. Psy., I, 1916, 494-507.
13. Chapman, J. Crosby & Toops, Herbert A. A Written Trade Test, Multiple Choice Method., Jour. of Appl. Psy., 3, 1919, 358-365.
14. Cody, Sherwin. Commercial Tests and How to Use Them. World Book Co., Yonkers, N. Y., 1919.
15. Committee on the Classification of Personnel in the Army, Various Members of. How the Army Uses Individual Differences in Experience. Psy. Bull., 15, 1918, 187-206.
16. Committee on the Classification of Personnel in the Army. The Personnel System of the United States Army. C. C. P. 339 & 340, Adj. Gen., U. S. A., Wash., D. C., 1919.
17. Frankfurter, Walter. Arbeitsversuche an der Schreibmaschine. Psychologische Arbeiten, 6, 1914.
18. Garnett, J. C. M. General Ability, Cleverness and Purpose. Bri. Jour. of Psy., 9, 1919, 345-366.
19. Haggerty, M. T. The Analysis of an Occupation. Manual Training & Vocational Education, 16, 1915, 472-479.
20. Hart, B. & Spearman, C. General Ability, its Existence and Nature. Bri. Jour. of Psy., 5, 1912, 51-84.
21. Heinetz, Wilhelm. Vorstudien über die psychologischen Arbeitsbedingungen des Maschinenschreibens. Zsch. für Angew. Psy., 13, 1918, 37-90.
22. Herberts, R. Zur Psychologie des Maschinenschreibens. Zsch. für Angew. Psy., 2, 1909, 551.
23. Heron, David. The Danger of Certain Formulas Suggested for Substitutes for the Correlation Coefficient. Biometrika, 8, 1911, 109-122.
24. Hollingworth, Harry L. Individual Differences Before, During and After Practice. Psy. Rev., 21, 1914, 1-8.



25. Hollingworth, Harry L. *Vocational Psychology*. D. Appleton & Co., 1918.
26. Hollingworth, Harry L. & Poffenberger, A. T. *Applied Psychology*, D. Appleton & Co., 1918.
27. Kelly, Truman Lee. *Educational Guidance*. Teachers' College Contributions to Education, 71, 1914.
28. Kelly, Truman Lee. Principles Underlying the Classification of Men. *Jour. of Appl. Psy.*, 3, 1919, 50-67.
29. Kelly, Truman Lee. Tables to Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations. *University of Texas Bulletin*, 27, 1916.
30. Kemble, William F. *Choosing Employees by Mental and Physical Tests*. Engineering Magazine Co., N. Y., 1917.
31. Kornhauser, A. W., Meine, F. & Ruml, B. Two Models Showing the Interrelation of Several Significant Correlation Variables. *Psy. Bull.*, 16, 1919, 230-234.
32. Lahy, J. M. Etude Experimentale de l'adaptation Psycho-physiologique aux Actes Volontaires Brefs et Intenses. *Jour. de Psy.*, 10, 1913, 220-236.
33. Link, Henry C. *Employment Psychology*. MacMillan Co., N. Y., 1919.
34. Lipmann, Otto & Stolzenberg, Otto. Methoden zur Auslese Hochwertiger Facharbeitern der Metallindustrie. *Zeit. für Angew. Psy.*, 16, 1920, 173-249.
35. Lough, James E. *Experimental Psychology in Vocational Guidance*. Proceedings of the Second National Conference of the National Vocational Guidance Association, N. Y., 1912.
36. Moore, Henry Ludwell. *Forecasting the Yield and Price of Cotton*. MacMillan Co., N. Y., 1917.
37. Munsterberg, Hugo. *Psychology and Industrial Efficiency*. Houghton Mifflin, N. Y., 1913.
38. Munsterberg, Hugo. *Vocation and Learning*. People's University, St. Louis, Mo., 1912.
39. National Association of Corporation Schools. *Bulletins issued Monthly*, N. Y., 1911.
40. National Vocational Guidance Association. *Proceedings of the National Conferences*, N. Y.
41. Ogburn, William F. & Goltra, Inez. How Women Vote. *Pol. Sci. Quar.*, 34, 1919, 413-433.
42. Pearson, Karl. On certain errors with regard to multiple correlation occasionally made by those who have not adequately studied the subject. *Biometrika*, *Miscellanea*, 10, 1914-1915, 81-87.
43. Pearson, Karl. On the General Theory of the Influence of Selection on Correlation and Variation. *Biometrika*, 8, 1911-1912, 437-443.
44. Pearson, Karl. On some novel properties of partial and multiple correlation coefficients in a universe of manifold characteristics. *Biometrika*, II, 1915-1917, 231-238.
45. Robinson, Edward S. The Analysis of Trade Ability. *Jour. of Appl. Psy.*, 3, 1919, 352-357.
46. Rugg, Harold O. *Statistical Methods Applied to Education*. Houghton Mifflin, N. Y., 1917.
47. Ruml, Beardsley. The Measurement of the Efficiency of Mental Tests. *Psy. Rev.*, 23, 1916, 501-507.
48. Ruml, Beardsley. The Reliability of Mental Tests in the Division of an Academic Group. *Psy. Rev. Mono.*, 24, 1917.
49. Sabatie, J. La Representation Commerciale; notions de psychologie professionnelle et appliquee. Dunod et Pinart, Paris, 1920.
50. Schneider, Herman. *Selecting Young Men for Particular Jobs*. American Machinist, April, 1913.

51. Soper, H. E. Probable Error of the bi-serial correlation coefficient. *Biometrika*, 10, 1915, 384.
52. Spearman, C. Correlation Calculated from Faulty Data. *Bri. Jour. of Psy.*, 3, 1910, 271-295.
53. Swift, Edgar James. The Acquisition of Skill in Typewriting. *Psy. Bull.*, 1, 1904, 295-305.
54. Swift, Edgar James. Memory of Skillful Movements. *Psy. Bull.*, 3, 1906, 185-187.
55. Swift, Edgar James & Schuyler, William. The Learning Process. *Psy. Bull.*, 4, 1907, 307-310.
56. Thompson, Godfrey H. The Proof or Disproof of General Ability. *Bri. Jour. of Psy.*, 9, 1919, 321-326.
57. Thompson, Godfrey H. The Hierarchy of Abilities. *Bri. Jour. of Psy.*, 9, 1919, 337-344.
58. Thurstone, L. L. The Learning Curve Equation. *Psy. Mono.*, vol. 26, 3, 1919.
59. Thurstone, L. L. Mental Tests for College Entrance. *Jour. of Educ. Psy.*, 10, 1919, 129-142.
60. Thurstone, L. L. A Scoring Method for Mental Tests. *Psy. Bull.*, 10, 1919, 235-239.
61. U. S. Bureau of Education. Pamphlets on Vocational Guidance, Wash., D. C.
62. U. S. Department of Labor. Bulletins issued Monthly, Wash., D. C.
63. Woodworth, Robert S. Combining the Results of Several Tests; a study in statistical method. *Psy. Rev.*, 19, 1912, 97-123.
64. Woodworth, Robert S. & Wells, Frederick L. Association Tests. *Psy. Mono.*, 13, 1911, No. 5, Whole No. 57.
65. Whipple, Guy M. Manual of Mental and Physical Tests. Warwick & York, Baltimore, Md., 1914.
66. Yule, G. Udny. Introduction to the Theory of Statistics. Chas. Griffin & Co., London, 1916.
67. Yule, G. Udny. On the Theory of Correlation. *Jour. of the Roy. Stat. Soc.*, London, 1897, 60, 812-854.





# ADENOIDS AND DISEASED TONSILS

## THEIR EFFECT ON GENERAL INTELLIGENCE

BY  
MARGARET COBB ROGERS, Ph.D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

No. 50

COLUMBIA UNIVERSITY CONTRIBUTIONS TO PHILOSOPHY AND PSYCHOLOGY

NEW YORK  
APRIL, 1922

226736<sup>c</sup>  
7. 11. 28

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Carey St., W. C.); Paris (16, rue de Condé)



## CONTENTS

### INTRODUCTION

Purpose of the Study. 5

### CHAPTER I. 7

Previous Literature Concerning the Relation of Nose and Throat Defects to Intelligence.

### CHAPTER II. 24

Method and Procedure.

1. A Statistical Study.
2. A Study of Improvement After Treatment.  
Selection of Cases.  
The Tests.

### CHAPTER III. 29

Discussion of the Results.

1. The Statistical Study.
2. The Study of Improvement after Operation.

### CHAPTER IV. 53

Measurement of Improvement after a Second Interval of Six Months.

### CHAPTER V. 68

Summary.

Conclusions.





## INTRODUCTION—PURPOSE

**D**URING the last decade or two there has been a growing interest among physicians in defects of the nose and throat. This interest has centered in part upon those two afflictions of childhood—adenoids and diseased tonsils,—or even tonsils that are merely enlarged. There is no doubt of the physical handicap borne by a child who is possessed of them. As a seat of inflammation, a source of infection, a hindrance to proper breathing,—in a multitude of ways they have seemed to deserve the verdict,—“Have them out.” Many physicians, to be sure, have cautioned against the wholesale removal of tonsils, saying that tonsils which are large in early childhood very commonly are absorbed at an early age.

But it is not my purpose to discuss the question of the efficacy of removing adenoids and tonsils. The aim of this study is, rather, to determine experimentally whether or not there exists any causal relation between defect in this respect and lowering of intelligence level. One hears statements made both pro and con by physicians and laymen, but there has been little experimental proof. It would seem to be rather useful for a physician to know in advance with how much probability of correctness he is speaking, when he advises a mother that the removal of adenoids and tonsils from the throat of her backward child will make him “bright.” The question in the present case, however, is broader than that of relation between these afflictions and mental defect. We are inquiring not merely whether adenoids and tonsils are causes of subnormality or dullness, but also whether they tend to lower the intelligence quotient in general however high it may be. Would the mentally normal child with adenoids and tonsils have been superior without them, and would the superior child have been still more superior? What is the relation between adenoids and tonsils, and intelligence?

The method employed in the present experiment would seem to give it value from the point of view of the clinical psychologist. With the present emphasis upon exactitude in mental testing, investigators have become interested in prob-

lem of the constancy of the I. Q. Adenoids and abnormal tonsils have been suggested as possible factors affecting this constancy. The results of the experiment should throw some light on the question.

It should be understood that this study is concerned with general intelligence, and not with the child's efficiency as a member of society. The latter question is much broader than the one we are investigating. It includes not only intelligence, but physical state, emotional make-up, volition: in short, the personality as a whole. Success in school work for example, depends upon all of these factors. For that reason, the results to be reported here, cannot be interpreted as applying to this broader conception. We cannot say at the end whether or not the physical defects under consideration affect the child's success as a member of society. We hope to be able, however, to determine their effect upon one element of that success, namely intelligence.

In presenting the results of this experiment, the writer is especially indebted to Professor R. S. Woodworth, under whose auspices the investigation was carried out, for his interest and advice; and to Dr. Leta S. Hollingworth for the suggestion of the problem, practical aid in obtaining subjects, and constant inspiration. She is indebted to the School of Education, Teachers' College, for the provision of operative treatment for the subjects; to Mr. Mark and his officers at Public School 64; and to Superintendent O'Brien of the Manhattan Eye, Ear and Throat Hospital. It must be said that by their hearty and generous cooperation they have contributed in a large measure to whatever value this study may possess.



# ADENOIDS AND DISEASED TONSILS: THEIR EFFECT UPON GENERAL INTELLIGENCE

## CHAPTER I.

### PREVIOUS LITERATURE

#### Concerning the Relation of Nose and Throat Defects to Intelligence

There are very few experimental studies of the relation between intelligence and the two defects considered here. There are a few statistical studies, and among earlier writers especially many statements of opinion on the matter. Characteristic of the latter is the following extract from an article in the Boston Medical and Surgical Journal, March, 1886.\*

“ . . . it is a fact that their intelligence may become weakened and their characters changed. They do not progress in their studies at school, are generally at the bottom of the class and remain in it longer than the prescribed time. . . . That the impairment of intellect and want of energy manifested by these children is real, and not merely in the expression of countenance, is made evident by watching these same children after the growths have been removed. To the gratification and astonishment of the parents and teachers, the children hitherto sluggish and dull of comprehension, now make rapid progress, and their comrades soon cease to make a laughing stock of them.”

The following quotation from an article by Irving Townsend, M. D., is in the same vein:†

“Aprosexia is the rather imposing term applied to the imperfect or arrested mental development attributed to this

---

\*F. Hooper, M. D., quoting from a paper by B. Frankel.

†Adenoid Growths of the Naso-pharynx. Read before the Homeopathic Medical Society of New York, February, 1895.

condition. This is denied by some authors, who claim that the dullness of comprehension and inattention are only apparent, and due only to defective hearing. A strong evidence of its reality lies in the fact that these children show most marvelous intellectual development after the removal of the growth, even in cases where deafness is not markedly improved."

A most enthusiastic denouncer of adenoids and abnormal tonsils is H. Addington Bruce. Concerning their direful effects upon the intelligence, and the magical results of their removal, he is continually reiterating:\*

"Often a surprising development of both mental and physical power follows the removal of adenoids. In one case reported by Professor Swift, a girl of fourteen grew three inches within six months after an operation for adenoids, and at the same time showed an improvement in her school work that contrasted strikingly with the dullness that preceded it. Another, three years younger, grew six inches in about five months, and from being a sad idler was transformed into an unexpectedly attractive and bright pupil. A boy of twelve, backward both mentally and physically, likewise lost his dullness and laziness within an astonishingly short time after the impediment had been removed."

And again:

"The boy or girl suffering from adenoids† is usually a mouth-breather because of the difficulty experienced in breathing through the nose. But mouth-breathing means difficult breathing, and this in turn means deficient oxidation of the tissues, with a resultant lowering of vital activities generally and of the activity of the brain in particular. Accordingly, the psychologist of today insists that every adenoid-afflicted child should be given prompt medical attention, with a view to correcting the vicious mouth-breathing habit, and thus aiding the child to gain a fair start in the development of mental and physical health."

The following extracts are quoted from Burgerstein's "Handbuch der Schulhygiene":

"Bresgen und Heymann machen endlich darauf aufmerksam, dass die Ursache der Kephalalgie haufig in der Behin-

\*H. Addington Bruce, *Psychology and Parenthood*, 1916.

†H. Addington Bruce in the *Century Magazine*, 1916—*The Mind of the Child*.

derung der Nasenatmung zu suchen ist, als Folgerscheinung von Verengerung der Nase bei ingen Baue des Knochengerüstes, Knochenkaries und Geschwulsten, Schwellungen der Schleimhäute, akuten Schnupfen, Verstopfungen der Highmorschöhle, Vergrosserung der Mandeln u. s. w." . . . . .

"Viele Kinder erscheinen schwachbegabt, ohne es zu sein, da bei denselben entweder nach behebung von Ohrenkrankheiten, nach Herstellung der freien *atmung oder Gebrauch einer entsprechenden Brille die scheinbare Geistesschwache schwindet.*"\*

Quotations like these, and equally unsupported by experimental evidence, might be multiplied indefinitely, especially if we look into the literature of a dozen years ago. Since they can have little authoritative value, I shall limit myself to two more specimens, one taken from the Psychological Clinic, 1916.†

"But when these physical defects (poor eyesight, defective hearing, adenoids, bad tonsils, etc.) are corrected so that the mind can function without any outcry from the physical body, these children recuperate mentally and often make greater progress than the so-called normal children in the regular grades."

The second is a quotation from Jelliffe and White, "Diseases of the Nervous System." Lee and Ferbiger, 1917, p. 903.

"An important group (of mental defects) is due to adenoid vegetations in the posterior pharynx. Under such conditions of ill health, development is impaired and does not proceed at a normal rate. With . . . . infected tonsils, which produce a constant toxemia, the child cannot be expected to proceed in his development with normal rapidity."

In the medical and psychological literature of the last few years, along with the growth of general discussion into the various phases of the operation itself, we find a general disinclination to take on faith the magic effect of adenectomy and tonsillectomy. This growth of critical spirit has shown itself in statistical investigations, and in studies of pedagogical and mental improvement after operation.

The statistical studies of physical defects in the schools reveal almost universally a positive relationship between

\*The italics are mine.

†Psych. Clinic, 1916, 10, 45-48. Anna Johnson. The Teacher in the Retarded School.



school retardation and possession of adenoids and diseased tonsils. One of these was conducted by Ayres for the Backward Children Investigation of the Russell Sage Foundation in New York City.\* The investigators examined the school records of 20,000 children from fifteen schools in Manhattan. Eight thousand of these had been examined by school physicians. The records of the physical examinations showed that 80 per cent of the children who were normal for their grade had physical defects while only about 75 per cent of the retarded children were physically defective.

This astonishing result was found upon retabulation of the data by ages, to be due to the fact that for each defect there is a gradual falling off in frequency from the age of six up to fifteen—eye-defect, only, excepted. Since the retarded children in each grade will be the older children in that grade, and since older children have fewer defects, the retarded children will show a smaller proportion of defect.

To overcome this difficulty, Ayres used an age basis instead of a grade basis in interpreting his results. Records of all the children at the ages of 10, 12, 13, and 14 were retabulated, a group of 3304 children, and rated as dull, normal or bright according to the grade in which they were found. The results were worked out in percentages of a group, and are shown in the following tables:

	Dull	Normal	Bright
Number of children examined .....	407	2588	309
Defects per child .....	1.65	1.30	1.07
Enlarged glands .....	20	13	6
Defective vision .....	24	25	29
Defective breathing .....	15	11	9
Defective teeth .....	42	40	34
<i>Hypertrophied tonsils</i> .....	<i>26</i>	<i>19</i>	<i>12</i>
<i>Adenoids</i> .....	<i>15</i>	<i>10</i>	<i>6</i>
Other Defects .....	21	11	11
Defective .....	75	73	68
Not defective .....	25	27	32

Average number of grades completed by pupils having no physical defects, compared with the number completed by those suffering from different defects:

\*Psych. Clinic, 1909, 3, 71-77. The Effect of Physical Defect on School Progress.

## 3304 Children, 10-14 years, grades 1-8

	Average grades completed	% lost
Children having no physical defects .....	4.94	
Children having enlarged glands .....	4.20	14.9
Children having defective vision .....	4.94	0
Children having defective breathing .....	4.58	7.2
Children having defective teeth .....	4.65	5.9
Children having hypertrophied tonsils .....	4.50	8.9
Children having adenoids .....	4.24	14.1
Children having other defects .....	4.52	8.5

Cornell reports several investigations in the Psychological Clinic, January and May, 1908. Three of these, in which children were rated on the basis of grades received in school work, are here combined to show the grades of normal children, "average" children, generally defective children, those possessing adenoids and tonsils, and the deaf.

No. of cases	Normal	Average	General Defective	Adenoids and Tonsils	Adenoids	Deaf
Allison		219				
9th St	64	84	21		8	
Claghorn	179	252	13			
Grade in language						
9th St.	72.9	70.5	63.3		60	
Claghorn	74.4	72.7	71.4			
Grade in Arith.						
9th St.	75.5	74	70		66.7	
Claghorn	72	70	65.1			
Grade in spelling						
9th St.	75.4	72.8	64.8		65	
Grade in geography						
Claghorn	76.6	76.5	76.2			
Average of grades						
Allison	75	74	72.6	72		67
9th St.	74.6	72.4	66		63.9	
Claghorn	74.3	73.1	70.8			

An additional investigation of four classes in the same grammar grade of the Claghorn School gives the following results:

	Class 1 Bright Children	Class 15	Class 9 Dull	Class 11 Dullest
Number of children .....	50	39	32	29
Normal .....	36	32	20	13
Defective .....	14	7	12	16
Percentage of normal .....	72	82	62.5	44.8

In the same article, Dr. Cornell gives the results of another study of Philadelphia schools, made in 1906. The study comprised a comparison of children exempt from examinations

on account of high standing, with those not exempt. The results follow:

	Exempt Children		Non-exempt Children	
	Normal	Defective	Normal	Defective
9th St. Primary .....	56	28	39	38
Rutledge School .....	87	35	75	34
Allison School .....	128	65	81	49
Camac School .....	183	71	103	75
Claghorn School .....	193	61	127	66
	<hr/>	<hr/>	<hr/>	<hr/>
Percentage Defective	647	260	425	262
	28.8		38.1	

When the four classes of bright and dull children were examined again, and the different sorts of defects compared for the groups, enlarged tonsils, adenoids, deafness, and nasal catarrh, were found to occur much more frequently among the two classes of duller children.

	Class 1 Bright Children	Class 15	Class 9 Dull	Class 11 Dullest
Number of children .....	50	39	32	29
Nose and throat conditions, number defective .....	6	4	9	9
<i>Tonsils</i> .....	3	4	3	3
<i>Adenoids</i> .....	2	1	5	6
Deaf .....	2		5	1
Catarrh .....			2	3
Percentage of children, nose and throat defects ..	12	10.2	28.1	31

During the same year, another examination along the same lines was conducted in the William McKinley Primary School,\* where a large number of dull children had been grouped in special classes.

None of these children were mentally defective, says Dr. Cornell, and only a few were really backward. The proportion of physical defect was found to be very large,—in 174 pupils, 188 physical defects (68 eye-strain, 40 nasal obstruction, 80 miscellaneous, 11 hypertrophied tonsils.) In a special class at the Wharton school, numbering 22 children, 14 of the children suffered from adenoids, associated in 3 cases with enlarged tonsils. Since no comparison is made with normal classes, this survey cannot be regarded as conclusive.

Wallin, in his book, "Mental Health of the School Child," discusses several other investigations of the relation of intelligence to physical defect. Only those studies in which were

\*Cornell, Psychological Clinic, 2, 1909.



included adenoids and tonsil conditions will be reviewed here. Those by Ayres and Cornell have been described above.

In Elmira, New York, says Wallin, "an investigation of repeaters in the second grade showed that 21 per cent of those who required three years and 40 per cent of those who required four years to complete the grade had adenoids, as against only 19 per cent of those who required only two years to do the grade."

Another study described by Dr. Wallin was made by Heilman in 1907 of 1000 Camden repeaters. The correlation between pedagogical retardation and percentage of defect in each group was as follows:

Defects	1 yr.	2 yr.	Retardation 3 yr. Per Cent	4 yr.	5 yr.
Health	16.5	21.3	28.0	19.0	37.5
Nutrition	13.4	8.9	17.2	20.2	17.5
Adenoids	6.2	7.3	8.1	9.6	7.5
Speech	5.2	5.1	4.2	10.5	20.0
Visual defects	15.5	15.9	18.2	22.8	22.8
Auditory	8.2	6.7	4.9	6.1	10.0

Burpitt\* describes an investigation of 400 children, 200 male and 200 female, considered by their teachers to be "dull and backward, but not to fall within the meaning of feeble-mindedness as given in the Mental Deficiency Act of 1913." The children were examined for physical defects and other abnormal conditions. The author says that in 36 per cent of the cases, the cause for backwardness was found to be "inherent dullness." (The basis for judgment of inherent dullness is not given.) Adenoids and tonsillar tissue were found in 18.75 per cent of the cases, and were "more prevalent than among the children of the area as a whole."

The degree of retardation, based upon the number of school standards below normal, was ascertained for pupils who suffered from various defects. The relative retardation was expressed by the fraction  $\frac{n}{A-5}$  where  $n$  = number of years retarded, and  $A$  = age. Eighteen per cent of the children were so retarded that the fraction was greater than 3-9.

\*H. R. Burpitt. Relative Degrees of Dulness and Backwardness in School Children and their Causation. *Journal of Mental Science*, 1916.

These were divided into two groups,—3-9 to 4-9 and 4-9 to 5-9. The results are given in the following table:

Causes	46 children 3-9 to 4-9	24 children 4-9 to 5-9
Inherent dullness only .....	8	3
Inherent dullness and one or more physical defects .....	7	2
Irregular attendance with one or more physical defects .....	9	6
Irregular attendance .....	2	3
<i>Adenoids only</i> .....	2	0

Turning to what the author calls single causes,—present in 170 cases out of the 400,—

Causes	151 children 1-9 to 3-9	19 children 3-9 to 6-9
Irregular attendance .....	51	6
<i>Adenoids</i> .....	24	2
Inherent dullness .....	59	11

The term “cause” seems to be rather loosely used in this study. The author says concerning this,

“Dealing with physical defects first, although they amount in the aggregate to 53 per cent (omitting defective speech, which is a secondary condition), in 10 percent only do they represent the whole cause. This is made up of those cases where the defect is of such intensity as to produce retardation in otherwise ordinary children, and of other cases of less intensity, but sufficient to weigh down the balance against those near the level of what we may call for convenience the lower limit of normal intelligence.” How he determines, without removing a defect, what the child’s intelligence would be without it he does not explain.

The following table compares the physical condition of two groups, one comprised of children examined in the regular routine examinations during the year 1912,—the other a group of retarded school children, given a special examination:—\*

	Group I	Group II
No. of children examined	287,456	1,541
No. with physical defects	206,720—71.9%	1,383—89.8%
No. of defects found	226,639	2,986

\*Transactions of the International Congress on School Hygiene, 1913, The Physical Condition of Retarded School Children.

Defect	No.	%	No.	%
Anaemia			335	23
Malnutrition	8,303	2.9	557	36.1
Defective vision	21,078	9.3	536	34.7
Defective hearing	1,206	0.5	47	3
Defective nasal breathing	21,931	7.6	316	20.4
Hypertrophied tonsils	30,021	10.4	297	19.2
Defective teeth	142,168	49.4	796	51.6
Pulmonary disease	335	0.1	47	3.0
Cardiac disease	1,597	0.5	35	2.3
Average No. of defects per child	1.1		2.5	

In an investigation of 3,587 exempt and 1,418 non-exempt children in the Philadelphia schools,\* Dr. Newmayer found the following percentages of defect:

Defect	Exempt Children		Non-exempt Children	
	No. Examined	%	No. Examined	%
Defective vision	371	10.0	171	12.0
Defective hearing	49	1.4	29	2.0
Defects of nose	54	1.5	21	1.5
Defects of throat	137	3.8	53	3.7
Orthopedic defects	25	.7	25	1.8
Mentally defective	6	.1	80	5.6
Skin diseases	918	26.0	423	30.0
Miscellaneous	214	6.0	128	9.0
Total	1,774	49.0	930	65.0

It is evident from the majority of these investigations that there is some relationship between physical defects and pedagogical retardation. But whether or not the relationship is a causal one, they do not indicate. Simple co-existence of two characteristics is not necessarily significant that one is cause of the other. Plainly, though, if the removal of a physical defect is followed by improvement in the school progress, it may be argued that the presence of the defect was a causal factor in the previous retardation. The method in the few following studies, which seems to be employed to a greater degree than formerly, consists of measurement of such improvement.

The Journal of Psycho-Asthenics, March and June, 1918, contains a paper on the "Results obtained from the Removal of Tonsils and Adenoids in the Feeble-minded," by Wm. J. G. Dawson, M. D. The author starts out rather discouragingly by regarding his hypothesis as an axiom. He says,

"It is a well-known fact that hypertrophy of the tonsils and presence of adenoids may produce more or less dullness

\*Ayres: "Laggards in Our Schools." 1909.



of the intellect in normal children. This is a result of the imperfect aeration of the blood which supplies the brain, on account of obstruction to respiration. In the feeble-minded, conditions are more or less similar."

One hundred and twelve cases in the Sonoma State Home, Eldridge, California, were operated on. Of these 6 are recorded as borderline, 39 as morons, 50 as imbeciles, and 17 as idiots. Adenoids were always removed when they were present. The results of the operation are as follows:

	Number before Operation	Number after Operation
Mouthbreathing	43	31
Eneuresis	33	32
Sore throats	70	2
Ear trouble	19	2
Change in voice		38 improved
Tonsillar tissue recurred in		5
General physical health		90 improved
		6 borderlines
		33 morons
		42 imbeciles
		9 idiots
Mental improvement from observation		27 improved
		4 borderlines
		15 morons
		7 imbeciles
		1 idiot

The inaccuracy of this investigation is evident. The mental improvement was measured by "observation," which is at best inexact, and susceptible to the influence of any expectation of improvement on the part of the observer. The degree of improvement is not mentioned, nor is the time interval allowed for the appearance of such improvement. There is no control group, and consequently, no way of knowing whether the improvement was due to the removal of the defect.

A similar, though rather more careful study is reported by Dr. Charles James Bloom in the New Orleans Medical and Surgical Journal for April, 1917. Dr. Bloom's experiment consisted of eighteen months' observation on the mental and physical state following the removal of adenoids and tonsils from one hundred and fourteen children. This number was later reduced to fifty-seven, because of the fact that a number failed to return. There was no selection, all the patients being taken as admitted.

The patient's physical and mental state was recorded at the time of admission. School reports were used as an index of intelligence. From this time on the patients were examined, weighed and measured at monthly intervals.

The ages of the children ranged from four to fourteen years, inclusive. Thirty-five per cent were under six years, and sixty-five per cent, therefore, over six. Twenty-nine were boys, twenty-eight girls.

Omitting a part of the study which though interesting has no bearing upon our problem, we turn to results in the way of mental status. There were fifty-seven cases, ten of whom were under the school limit. Of the remaining forty-seven, seven sent in no report. In four, or ten percent of the forty remaining, there was no progress. In thirty-six, or ninety per cent, appreciable progress was reported. One of the four unimproved cases was syphilitic, the other, the author says was a moron.

Quotation of the teachers' reports will be of interest.

"'Some improvement.' 'Better work than previous year.' 'More effort displayed.' 'Improved wonderfully.' 'Improvement first term, not so much second.' 'Before removal, not transferred; after removal transferred.' 'Very much improved, both mentally and physically.' 'Has made progress.' 'Remarkable improvement.' 'Not transferred before removal, but after.' 'More attentive.' 'A very small but gradual improvement.' 'Am happy to tell you that he is studying more since tonsils and adenoids were removed.' 'Greatly improved.' 'Attention better.' 'More concentration.'"

In this experiment like the preceding, the judges are liable to the effect of expectation of improvement. Although the reports are more explicit, they are still couched in general terms, and not commensurable. Some reports refer to intelligence and some to pedagogical standing. There is no control group.

On the basis of these results, the author concludes:

"Children exhibiting some alternatives in the normal histology of tonsils and adenoids, give marked evidences of mental impairment." This seems to be a rather sweeping statement in consideration of the number of intellectually superior children who suffer from adenoids and diseased tonsils.

Another investigation was made by Dr. Cornell in the Phil-

adelphia schools,\* where seventy more or less retarded pupils in grades one to four were operated on for adenoids. According to the teachers' reports—

30 per cent improved considerably.

40 per cent improved.

25 per cent did not improve.

1.6 per cent deteriorated.

3.0 per cent deteriorated considerably.

Of those who had two chances of promotion,

6.3 per cent were promoted twice.

16.0 per cent failed twice.

33.3 per cent were promoted once.

33.3 per cent failed once.

With one opportunity,

11.0 per cent were promoted.

31.7 per cent failed.

"The promotion record was thus decidedly poor. It is possible, however, that the time for promotion came before the orthogenic effects of the operations had become effective."

The same criticisms may be brought against this investigation as were mentioned in connection with the preceding ones. Teachers' estimates of improvement, especially when such improvement is expected, and without means of measuring it objectively, are necessarily inaccurate. Again there is no control group. Of even less value are the results of an investigation in New York City by Cronin, where, out of eighty-seven cases operated on for enlarged tonsils and adenoids, "many advanced three grades during the rest of the school year, and only three lost time."

An interesting study is one that is described by John C. Simpson, M. D., in the *Journal of the American Medical Association*, April 1, 1916.

Dr. Simpson's results are based on a study of 571 boys of Girard College who had been operated on for adenoids and tonsils. Improvement was studied along several different lines, among them scholastic ability. For this part of the study, 45 were chosen alphabetically, 3 from each section. The only selection was for boys who were operated on long enough after coming to school to give an idea of scholastic ability; and long enough before the present study to permit a judgment as to

---

\*Wallin: "Mental Health of the School Child." 1914.



their improvement. Monthly averages were taken of each boy up to the time of the operation and from then to the time of this study. They were based on an average of 100 per cent. As a control group, there were chosen 45 boys who had had no operation, and who lived and worked under the same conditions. They also were taken alphabetically, 3 from each section.

The general average of the operative cases at the first measurement was 74.04. Of these 25, or 55.5 per cent gave an average increase in monthly standing of 4.45 after operation, while the remaining 20, or 44.5 per cent suffered a decrease of 6.09.

The average of the boys in the control group was 74.21 and for the first group after operation 74.06. "It is interesting to note," says the author, "that the standing of slightly more than half of those operated on was improved, but when compared with those not operated on, no difference is seen."

In a similar study of younger boys who had undergone the operation on entering college, and who had since had a year's study (again a group of 45), the general average was 76.61. Compared to 45 in the same classes not operated on, who had an average of 74.56, the operative group is very slightly superior, 2.05 points.

Another study of pedagogical improvement, and a valuable contribution, is that reported by A. H. McPhail in *Pedagogical Seminary* for June, 1920, entitled "Adenoids and Tonsils; a Study showing how the Removal of Enlarged or Diseased Tonsils affects a Child's Work in School."

"The children studied were pupils in the Adams and Cranch Schools. Only cases were considered where there was a record of ten school months before the date of the operation, and where there was a record for at least ten month after the operation. There were thirty-one cases in all.

"School records were looked up for the ten school months preceding operation, and for each school month subsequent—up to the date of leaving school, or in the case of children still in school, up to the date of the study. There were thus longer school records for some than for others.

"The history of each case was divided into periods of ten school months each. Eighteen cases had a record of twenty months after operation and eleven cases of thirty months.

"Comparing the first period after the operation with the period before, it is found that only about one-third showed improvement, and a little over half were doing poorer work. By comparing an average of *all* work done subsequent to the operation with what was done before, it became evident that improvement in school work is not often observed until after a year from the date of the operation."

In the cases that had records for twenty months, 16 show that better work was done in the second period after the operation than in the period before.

TABLE A

Table showing where improvement begins.

	<i>Per Cent of cases showing improvement</i>		
	Period 1	Period 2	Period 3
Based on 31 cases	32.2		
Based on 18 cases	33.3	66.6	
Based on 11 cases	36.3	63.5	100
Based on all groups	33.9	65.	100

These cases were compared with a control group chosen at random. They comprised a total of 100 children who had records for four consecutive school years.

TABLE B

Table showing how time of improvement of "operated cases" compares with improvement among children at random.

	Period 1	Period 2	Period 3
A—Per cent of 31 cases (operated upon)	32.2		
Per cent of 100 children at random	42.		
B—Per cent of 18 children—20 months (operated upon)	33.3	66.6	
Per cent of 100 children at random	42.	41.	
C—Per cent of 11 cases for 30 months (operated upon)	36.3	63.5	100
Per cent of 100 children at random	42.	41.	41

". . . . Immediately after operation, there seems to be a dropping off in the quality of school work done," but thereafter a marked improvement while the random group shows a comparatively static percentage of improvement from year to year. The conclusion of the author is:

"Here seems good reason to believe that the removal of diseased tonsils and adenoids is a factor in beneficially influencing the mental life of the school child. Not only is the health impaired by failure to remove these diseased parts but the mental life and activity of the child as well."

It is conceivable that pedagogical retardation might exist without any defect of intelligence. The physical effects of adenoids and tonsils might produce a tendency to fatigue, an emotional instability and consequent lack in attention, which would seriously influence the quality of school work, even though the child were of normal or superior intelligence. The relation of physical defects to intelligence has been investigated experimentally by a method which will be employed to some extent in the present investigation. In the two studies to which I refer psychological tests, rather than school standings were used as a basis for judging the intelligence. In each the effects of treatment were measured, and in one, a control group makes possible a more accurate interpretation of results.

The first of these investigations is described by Wallin.\* It was "an attempt to determine by controlled, objective, mental measures the influence of hygiene and operative dental treatment upon the intellectual efficiency and working capacity of a squad of twenty-seven public school children in Marion School, Cleveland, Ohio (ten boys and seventeen girls), all of whom were handicapped to a considerable degree with diseased dentures or gums, and an insanitary oral cavity." The experiment extended over one year, from May, 1910, to May, 1911. The treatment included corrective work upon the teeth and mouth, and also instruction in oral hygiene, and follow-up work by an employed nurse. Five series of psychological tests were given at stated intervals during the course of the experiment. They included tests of immediate recall, spontaneous and controlled association (opposites), adding, and attention-perception (cancellation). There were six sets of each test, numbered from one to six, of equal difficulty, and given under uniform conditions. Tests 1 and 2 were given before the treatment began, and the average was taken as the "initial efficiency." The last four, or the last two, were averaged to represent the pupils' "terminal efficiency."

---

\*Wallin: "Mental Health of the School Child." 1914.



The results show the following influence of dental treatment upon the working efficiency of the pupils.

1. The indices of improvement are about the same for boys and girls.

2. Improvement was about the same for older and younger pupils.

3. There were great individual differences in initial proficiency and in improvement.

4. Improvement in one test does not presuppose improvement in another.

5. There is a decided gain in every test, "and not only are the gains decidedly more frequent than the losses but the largest gains are invariably emphatically larger than the largest losses."

6. The average gains in the tests were:

Memory, 19 per cent with 8 losses and 19 gains.

Spontaneous association, 42 per cent with 2 losses and 25 gains,

Addition, 35 per cent with 1 loss and 26 gains.

Controlled association, 29 per cent with 0 losses.

Perception-attention, 69 per cent with 0 losses.

Average gain for all tests, 57 per cent.

Unfortunately, Wallin was unable to form a control group, so that it is impossible to estimate accurately how much of this gain is due to the treatment of the defect, and how much to other causes, such as growth, etc. "But," the writer adds, "if we concede that one-half of the gain—and that is, I believe, a sufficiently liberal concession—is due to a number of extrinsic factors, such as familiarity, practice and increased maturity, the gain solely attributable to the heightened mentation resulting from the physical improvement of the pupils would still be very considerable. There is corroborative evidence to show that there was a general improvement in the mental functioning of these pupils. This evidence is supplied by the examination of the pedagogical record of scholarship, attendance and deportment. Most of the members of this experiment squad were laggards, and repeaters, pedagogically retarded in their school work from one to four years, but during the experiment year only one pupil failed of promotion,

while six did thirty-eight weeks of work in twenty-four weeks, and one boy finished two years of work within the experimental year."

The second investigation was equally careful in its method. It was pursued by the Rockefeller Foundation, under the direction of E. K. Strong, with the purpose of examining the "Effects of Hookworm Disease on the Mental and Physical Development of Children."

The children were divided into five groups and tested at intervals of three and one-half months. The tests used were opposites, calculation, logical memory, memory span, handwriting, formboard, and Binet-Simon. After the first test-series was given, the five groups were divided into sub-groups on the basis of this initial performance, so that the improvement was compared only for those sub-groups in which this was equal.

The improvement of Group A—uninfected children—proved to be greatest, and was taken as 100 per cent. On this basis, Group B—infected children not treated—showed the least improvement,—only 34 per cent. Group C—children completely cured of infection—improved 60 per cent. Group D—severely infected children, treated but not completely cured—improved 38 per cent, and Group Du—an older sub-group of D—improved 9 per cent as much as the normal children, and much less than the untreated younger children. Dr. Strong reaches the following conclusion:

"The figures show, then, that hookworm disease unmistakably affects mental development. Treatment alleviates this condition to some extent but it does not, immediately, at least, permit the child to gain as he would if he had not had the disease. And the figures apparently further show that prolonged infection may produce prolonged effects upon mentality,—effects from which the individual may never recover."

## CHAPTER II.

### METHOD AND PROCEDURE

The following investigation was carried on during the year and a half from October, 1919, to April, 1921. The subjects were pupils at Public School 64, Manhattan, or patients at the Manhattan Eye, Ear and Throat Hospital. All were boys, between the ages of six and fourteen years. The testing in the study of improvement was done by the investigator, assisted by three other examiners, all competent and experienced in the technique of giving psychological tests.

#### A Statistical Study

In addition to the more lengthy experiment, a statistical study was made, comparing the intelligence levels of two groups of children, the one selected for the presence of tonsils, the other for freedom from them. These two groups were obtained from a large group of 530 children whose I. Q.'s were gained from the records of Public School 64, where, so far as possible, all children are tested upon entering school. We had, therefore, a group unselected for intelligence level.

All the children for whom we had I. Q.'s were examined by the school nurse or physician. On the basis of this examination the two groups were selected. The tonsil group consisted of those cases which in the opinion of the nurse or doctor, were pronounced enough to deserve treatment. The normal group was composed of those who were not defective, or in whom the defect was so slight as not to demand treatment. The two groups were arranged each in a surface of distribution according to the I. Q.'s of the members. The results of the distribution appear in Table I, and in Figs. I. and II.

#### A Study of Improvement After Treatment

The method employed here is based on the hypothesis that if a physical defect is the cause of retardation in mental or physical development, removal of the cause will tend to lessen



the retardation. In other words, if a child's working efficiency is lowered by the effects of adenoids and bad tonsils, their removal should, unless such lowering be permanent, be followed after a reasonable time by an improvement. But improvement in efficiency, following the removal of adenoids and tonsils proves nothing unless we shall compare it with the change in efficiency of a control group, whose members have not been operated on, and who thus still suffer from the effects of the growths.

### Selection of Cases

The selection of the children for the experiment was effected in the following manner. The teachers at Public School 64 were asked to report any cases which had come to their notice, as being seriously afflicted with adenoids and diseased tonsils. In this way a fairly large group was obtained. The parents of the children were visited with the purpose of obtaining permission for examination and operation at the Post Graduate Hospital. It was fairly easy to obtain permission to have the children examined. They were taken in groups of four or five to the clinic, the experimenter attending in person every examination in order to learn from the doctors the degree of the defect. As a result of this method, we discarded all those cases where there was any doubt as to the serious nature of the defect.

From the large group examined, we were finally successful in securing operative treatment for 10 children. Discarding the cases where defect was slight, there remained a number of children who, for one reason or another, did not undergo operation. In some instances the parents refused their permission, in some they failed to keep appointments, in one or two there was sickness in the family, and in a number the hospital was overcrowded and could not receive the children. All members of this group were examined,—to the number of fifty-six, and from them the control group was finally selected.

Since we were unable to secure a large test group from Public School 64, the experiment was continued at the Manhattan Eye, Ear and Throat Hospital where opportunity was given for testing the children after they had been admitted for operation. In order to be sure that in each case the defect was sufficiently pronounced to render decisive the results of the experiment, each child's card was examined. Only those

children were included who were undergoing operation for both adenoids and tonsils.

It may be here remarked that mental tests were given to these children on the morning of operation, and in some cases only a short time before it. The possibility suggests itself, therefore, that the results of the tests may have been influenced by excitement or fright on the part of the patients. Actually, however, this did not seem to be the case. The children were perfectly cheerful and showed no signs of nervousness. The tests were given in a waiting room with the door closed so that any disturbing sights and sounds were eliminated. The possible lowering of the performance by the causes mentioned would tend to exaggerate the improvement shown by the re-tests, so that in the light of the results, it will be seen that they could have had little effect.

The test group, then, was composed of forty members; ten from Public school 64, who received operation at the Post Graduate Hospital, and the remaining thirty from various schools throughout the city, patients at the Manhattan Eye, Ear and Throat Hospital. The control group of forty was selected as previously described, and the pairs were arranged so as to have the ages of the members of one pair as nearly as possible the same.

### The Tests

Since the main interest of this investigation lies with intellectual development, two tests of intelligence were given: namely, Terman's revision of the Binet test, and Healy's Picture Completion Test, number II. The starred Terman was always used, since it was necessary to economize time.

It was expected that improvement in general health would probably follow the removal of the defects. This physical gain should come to light in increased height and weight. In every case, therefore, height and weight were measured.

It is conceivable that adenoids and tonsils might have no effect upon general intelligence, and yet might cause a noticeable pedagogical retardation, simply as a result of the child's physical handicap, tendency to fatigue and consequent defect in attention or sustained effort. In order to gain some measure of this physical factor, strength of grip and speed in tapping were found. An effort was made, also, to obtain teach-

ers' estimates of the pedagogical rankings, but this was for the most part unsuccessful, since in many cases teachers misunderstood directions, and in others the tests were made too soon after the opening of school for any such estimates to be possible.

The tests described above were given before the operation to each child in the test group, allowing as short an interval as possible between test and operation. In the case of the Manhattan Hospital children, test and operation fell on the same day. In no case did the interval exceed ten days. The members of the control group were tested, each one within a week of his partner.\*

Six months after his first test, each child was re-tested, whenever possible. Since some children had dropped out of the groups for one reason or another, the final number in each group was twenty-eight. It was necessary to rearrange the control cases somewhat in order to fill in spaces left vacant by those who were lost. In this rearrangement, the effort was made, 1. to pair cases whose ages were approximately the same; 2. to pair cases whose first tests were dated fairly close together. Since all the children were tested and re-tested under approximately the same conditions, this rearrangement will probably not greatly influence the results. The tests were always given in the same order.

The following table shows a list of the two groups, as originally paired, and as finally rearranged, with dates of tests and retests. Dates of operation are given for the first group.

---

\*In a few cases where the operation was postponed after the test had been given, the child and his control were retested just previous to the operation. Since both cases were retested, practice effect is of no great importance.



Test I	Test Case Op.	Test II	Original Control		Test II	Final Control	
			Test I			Test I	Test II
JB	10-15-19	lost	SS	10-15-19	4-15-20	SS	10-15-19
LL	10-15-19	4-15-20	LJ	10-15-19	lost	MG	4-30-20
HK	10-20-19	4-30-20				AA	5-17-20
MS	11- 6-19	5-17-20				SD	6-11-20
GF	12-26-19	6-11-20				NF	5-14-20
RJ	12-16-19	12-30-19				ML	6- 9-20
JJ	12-16-19	12-30-19				LP	7-15-20
AG	1-15-20	1-16-20				AL	8- 2-20
IK	2-14-20	2-16-20		control removed		JF	2-11-20
HG	2-10-20	2-11-20				JF	8- 3-20
AC	2-11-20	2-12-20					8- 3-20
CL	2-26-20	3- 1-20		control removed		PG	2-26-20
MR	2-26-20	3- 1-20					8- 3-20
SR	2-26-20	moved		control removed		SK	9-24-20
IK	3-17-20	moved		control removed			
AO	3- 8-20	9-20-20		control removed			
RB	3- 8-20	moved		control removed			
DT	3- 8-20	mastoid		control removed			
AL	3- 9-20	moved					
JD	3- 9-20	9-23-20					
LS	3- 9-20	9-25-20					
JB	3-12-20	moved		control removed			
HS	3-13-20	9-21-20	JM	3-13-20	lost	DD	9-16-20
AM	3-13-20	9-20-20	SS	3-22-20	wrong boy	KS	9-24-20
SO	3-18-20	9-22-20		(adenoids		MR	9-15-20
IF	3-18-20	9-23-20	LC	3-22-20	(removed	MA	10- 1-20
AD	3-19-20	9-20-20	IB	3-23-20	9-24-20	PK	9-23-20
JR	3-19-20	moved	MA	3-13-20	9-23-20	IB	9-21-20
JN	3-20-20	3-20-20				LF	10- 1-20
HS	3-20-20	9-21-20				SB	9-21-20
II	3-26-20	9-24-20				BF	10- 1-20
UF	3-27-20	9-29-20				LF	10- 1-20
SM	3-27-20	9-30-20				LG	10- 1-20
AM	3-29-20	9-29-20				4- 6-20	10- 1-20
CK	3-29-20	9-29-20				BG	10- 1-20
FB	3-30-20	9-29-20				NF	10- 1-20
AA	3-30-20	9-23-20				JF	10- 1-20
LS	3-31-20	moved		control removed			9-30-20
FT	3-31-20	9-28-20				MA	
LP	4- 1-20	moved	HH	4- 6-20	10- 1-20		

## CHAPTER III.

### DISCUSSION OF THE RESULTS

#### Statistical Study

The statistical study compared two groups of cases in respect to I. Q. These groups were selected from one large group, on the basis of presence or absence of tonsillar defect. The tonsil group was composed of 236 cases, and the normal group, of 294. The distribution of the two groups according to intelligence is set forth in Table I, and in Figs. I and II.

TABLE I

I. Q.	Tonsil Group		Normal Group	
	No. of Cases	Per cent of Cases	No. of Cases	Per cent of Cases
40- 50	2	.8	0	0
50- 60	1	.4	2	.7
60- 70	7	2.9	4	1.4
70- 80	21	8.9	29	9.8
80- 90	45	19.0	52	17.7
90-100	80	33.9	107	36.4
100-110	55	23.3	67	22.8
110-120	17	7.2	24	8.1
120-130	6	2.5	9	3.0
130-140	2	.8	0	0
140-150	1	.4	0	0
Average	94.9		95.4	
Median	95.3		95.6	
Q	8.705		8.27	
$\sigma$	14.4		12.2	

From these it is evident that the two groups are practically equal in intelligence. The average I. Q. for the normal group is 95.4, as compared with 94.9 for the tonsil group. The medians are equally close,—95.6 in the normal group and 95.3 with the tonsil cases. The difference in variability is negligible, Q being 8.705 and  $\sigma$  14.4 in the tonsil group, while in the normal Q is 8.27 and  $\sigma$  12.2. The two cases with the lowest I. Q.'s were tonsil cases, but the three highest I. Q.'s also belong in this group.

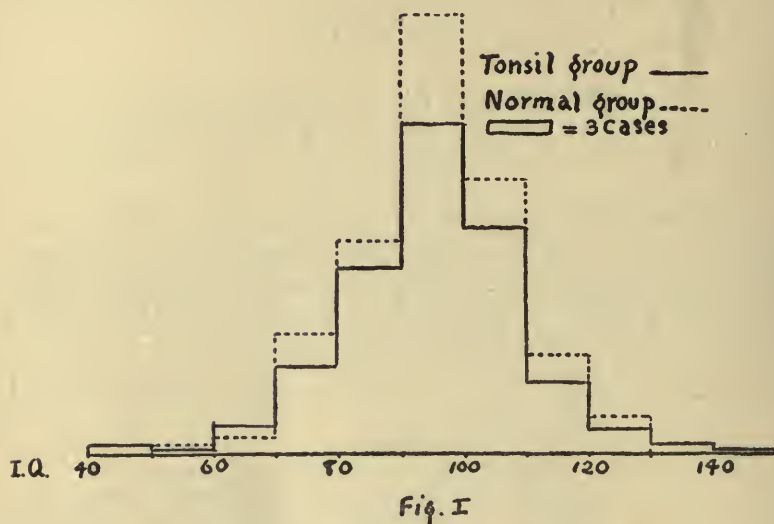


Fig. 1. Distribution of I. Q.'s. Number of cases.

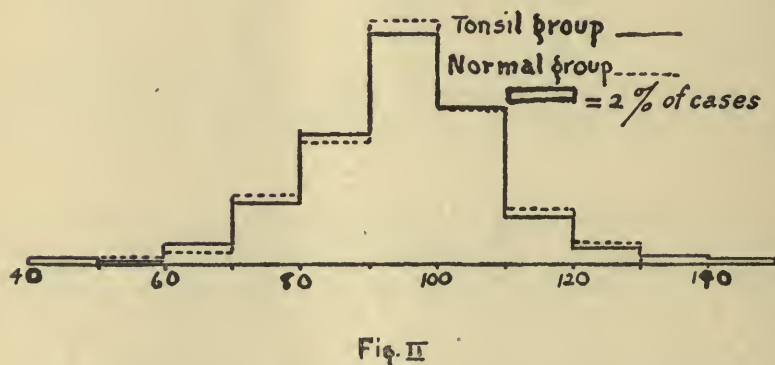


Fig. 2. Distribution of I. Q.'s by percentage of total number of cases in the group.



If the frequencies are expressed in terms of per cent of the total number of cases in the group, the two may be compared further. The following details are noticeable.

I.Q.	Per cent of Tonsil Group	Per Cent of Normal Group
Below 70	4.1	2.1
Below 90	32.0	29.6
Above 110	10.9	11.1
Above 120	3.7	3.0
Above 130	1.2	0

In other words, in the percentage of cases below normal intelligence, the tonsil group exceeds by 2.4 per cent. The percentage of defective cases is also slightly greater in the tonsil group—the difference here being 2 per cent. The normal group has a negligible predominance of bright cases,—only two-tenths of one per cent difference, while with the very superior cases, the tonsil group again exceeds,—by 1.2 per cent. The per cent of the tonsil group which reaches or exceeds the median of the normal is 49 per cent.

These figures seem to indicate remarkable similarity between the two groups considered. The two distributions are almost identical. While the slight predominance of cases below normal mentality in the tonsil group may indicate a very feeble tendency toward coincidence of tonsillar defect and mental dullness, it does not seem large enough to be at all significant. This is especially true when we consider that the tonsil group exceeds in superior children. If we allow the preceding contention of coincidence between dullness and tonsils, must we not argue here in the same manner for a tendency toward coincidence of superiority and tonsils?

The chief source of error in this part of the study is the fact that the throat examinations were not conducted by the same person throughout the investigation. For his reason there must have been some slight disagreement as to what should constitute a reportable case. In the event, then, of a positive relationship between tonsil defect and lowering of the intelligence quotient, placement of normal tonsils in the "tonsil" group, and of diseased tonsils in the "normal" group would raise the first, and lower the second, thus tending to conceal the difference between the two. On the other hand, the cases where disagreement would occur would naturally be those of slighter defect, in which the intellectual retardation would be less likely to occur, so that the result would

probably be merely an increased height at the overlapping portion of the curves, with no change at the ends.

In any case, the two examiners had worked together previously, so that each must have been somewhat familiar with the opinions of the other. They were aware, also, that pronounced tonsillar defect was what we were attempting to detect. However this may be, there must always be some disagreement in diagnosis. When this is allowed for, the results of the investigation may be taken for what they are worth. Contrary to expectation, there seems to be very little difference in intelligence between a group of children whose throats are normal, and one in whom the tonsils are diseased or badly enlarged.

### STUDY OF IMPROVEMENT AFTER OPERATION

The complete results of the tests and retests are collected in Table II, where each control case is listed immediately below its respective test case, and where age, height, weight, grip, tapping rate, I. Q., and score in Healy Picture Completion are shown. From these data the more detailed observations have been made. The improvement of each child in the various tests has been computed, and a comparison drawn between the two groups. As we have previously stated, any improvement shown by the test group in excess of that of the control group, may be looked upon as significant.

Let us consider first the improvement of the children in general health, as shown by height and weight. In Tables III and IV we have tabulated the results, in such shape as to permit of comparison. An inspection of these tables will establish the fact that after a six months' interval, the test group shows, in respect to height and weight, a very slight gain over the control group. In weight, the average of the amounts by which the test group gains exceed the control group gains is 1.37 lbs., and in height, only .16 inches. The medians of these amounts are 1.2 lbs. and .2 inches respectively. Comparing the improvements for the two groups, we find that in the case of the weights, the smallest gain (a loss of 1.2 lbs.) occurs in the control group, while the largest gain (10.7 lbs.) is in the test group.

TABLE II. RESULTS OF TESTS

Blank spaces indicate where tests were omitted for one reason or another

N	Age		Weight lbs.		Height in.		Grip, Kg. best hand	
	1	2	1	2	1	2	1	2
1	7- 7	8- 1	50.4	54.2	46	47.6	13	12
1C	8- 1	8- 7	53.5	57.2	46.4	47.8	11	13
2	6- 9	7- 3	40.9	42.9	42.6	41.1	9	9
2C	7- 1	7- 7	52.3	57.4	45.2	47	10	12
3	8-	8- 6	55	59.5	47	48.4	12.8	14.5
3C	9- 9	10- 3	61.5	62.9	51.7	52.9	14	15
4	8-10	9- 4	51.1	54.2	47.5	49.2	9	—
4C	9-10	10- 4	49.4	51	48.9	52	9.5	—
5	6- 1	6- 7	45	47	44.9	45.2	11	—
5C	8- 2	8- 8	56.2	57	46.6	48.1	12	—
6	5- 2	5- 8	43.8	44.5	43.1	43.9	8	—
6C	7- 1	7- 7	50.6	52.5	45.4	47.3	10.5	—
7	6- 7	7- 1	39.9	41	42.9	44.8	7	6.5
7C	6- 7	7- 1	38.4	38.7	41.9	43.2	9	10
8	8- 6	9-	60.8	63.3	50.8	51.8	10	—
8C	8- 5	8-11	45.4	52.1	46.8	47.6	15	16
9	9- 4	9-10	50.6	53.2	48.1	49.4	10.5	13
9C	9- 6	10-	59.8	61.4	51.9	55.2	16.5	21
10	6- 7	7- 1	48.9	51.4	46.1	47.7	12.5	11
10C	7-	7- 6	47.1	47.5	45.6	47.2	10	15
11	6- 7	7-	47.8	47.5	45.8	47.7	11	15
11C	6- 8	7- 1	41.6	42.5	43.6	44.9	11.5	11.5
12	7- 8	8- 2	48	52.5	44.8	—	14	—
12C	7- 1	7- 8	41	44.5	41.5	43.3	6	4.5
13	13- 3	13-10	90	98	61.5	65	26.5	28.5
13C	14- 6	15-	74.7	76.8	56.8	57.8	22	23
14	11- 9	12- 4	56	62	51	51.6	16	15
14C	11-10	12- 4	81.9	86	57.9	58.3	22	24
15	10- 3	10-10	57.5	—	51.1	—	15.5	—
15C	10- 1	10- 7	67.2	70.3	50.1	51	15	15.5
16	10- 9	11- 3	56	57	51.6	52.3	19	17.5
16C	10- 9	11- 3	51.2	50	48.7	49.5	10	10
17	8- 1	8- 7	57	—	48.7	—	14	—
17C	7-10	8- 4	45.3	—	44.8	—	10	8.5
18	7- 2	7- 8	58.2	—	47.3	—	11	—
18C	6-11	7- 5	45.3	47	46.7	47.1	8	6.5
19	11- 4	11-10	90	96.3	57.7	59	22	21
19C	7-11	8- 5	52.4	54.4	46.7	47.2	15	12



TABLE II. RESULTS OF TESTS (Continued)

Blank spaces indicate where tests were omitted for one reason or another

N	Age		Weight lbs.		Height in.		Grip, Kg. best hand	
20	7- 1	7- 7	44.2	—	47.2	—	11.	—
20C	7- 3	7-10	61.3	66	49.6	55	15	12.5
21	11-	11- 6	70.7	76.5	54.1	—	16.5	16.5
21C	10- 1	10- 7	62.4	67	49.6	50.4	19	15
22	10- 9	11- 3	73.3	—	53	56.4	18	22.5
22C	11- 7	12- 1	70.7	80.5	56.8	58.1	19.5	21.5
23	8- 7	9- 1	51.7	—	47.8	—	11.5	15.5
23C	8-11	9- 4	64.1	66.5	51.4	53.1	14.5	14
24	9- 8	10- 2	58.5	62.5	51	—	19	20
24C	10- 2	10- 8	60	61	50.1	51.5	15	15
25	10- 1	10- 7	55.5	59.5	50	50.8	14	—
25C	10-10	11- 4	63.3	63.8	50.2	50.9	12.5	21.5
26	9- 8	10- 2	63.8	74.5	51.6	54.3	14	—
26C	10- 4	10- 9	64.2	67	51.4	52.3	20	16.5
27	6- 7	7- 1	43.7	—	45.4	—	9	6
27C	6- 3	6- 9	41	44	44.6	45.4	8	9
28	12-11	13- 5	71.3	75.5	54.9	55.8	23.5	21
28C	13- 8	14- 2	74.2	79.8	53.4	54.5	21	—

TABLE II. RESULTS (Continued)

Blank spaces indicate where tests were omitted for one reason or another

N	Tapping, ½ min. best hand		I. Q.		Healy, Score	
	1	2	1	2	1	2
1	135	120	82	83	—25	— 2
1C	106	115	80	76	—50	—16
2	105	112	107	114	28.5	30
2C	152	114	91	96	3	—11
3	136	139	94	91	21.5	22.5
3C	135	129	82	85	17	19
4	103	—	96	96	8.5	—
4C	109	—	83	85	33	—
5	110	—	95	99	—25	—
5C	156	—	114	117	40.5	—
6	110	—	95	101	—33	—
6C	126	—	88	89	—32	—
7	125	113	91	99	6	—28
7C	105	—	95	99	4.5	27.5
8	113	110	91	86	32.5	—
8C	131	101	98	104	4	23

TABLE II. RESULTS (Continued)

Blank spaces indicate where tests were omitted for one reason or another

N	Tapping, $\frac{1}{2}$ min. best hand		I. Q.		Healy, Score	
9	149	135	83	93	3.5	10.5
9C	144	150	87	90	34	55
10	68+ 74	88+ 82	110	109	-12	6.5
10C	70+ 54	135+109	104	100	27	65
11	125+ 90	98+ 87	103	100	- 8	6
11C	155+125	101+107	101	102	-29	- 3.5
12	98+ 69	—	98	95	20	21
12C	102	84	98	101	-10	-12
13	160+165	142+134	70	78	43	42
13C	150+109	122+ 94	66	64	- 1.5	30.5
14	190+172	138+130	96	107	12.5	48.5
14C	175+152	175+164	140	137	- 5	25.5
15	172+167	170+156	97	94	7	25
15C	140+115	137+115	78	79	1	42.5
16	145+131	—	65	73	49	47.5
16C	145+ 99	135+135	74	82	30	37
17	90+ 89	150+100	71	77	29.5	12
17C	125+116	121+ 97	96	99	1.5	15
18	133+115	135+111	98	98	-13.5	-12
18C	100+ 99	84+ 74	90	94	-32	-28
19	168+136	—	96	101	57.5	49
19C	100+115	118+ 92	98	98	-22	-11
20	105+115	110+ 93	106	102	0	-11
20C	150+120	155+149	118	131	30	35
21	152+111	132+125	64	67	20	32
21C	140+136	138+110	86	97	70.5	58.5
22	164+148	183+141	91	100	48.5	43.5
22C	120+116	157+127	63	62	34.5	33.5
23	150+119	141+136	85	94	49.5	68
23C	122+115	140+110	81	96	4	25
24	157+136	142+126	131	124	54.5	63
24C	155+135	155+100	89	92	31.5	59.5
25	140+127	150+119	77	76	8	25
25C	148+134	151+135	145	137	29.5	29
26	137+113	138+117	80	76	22.5	7
26C	125+105	125+ 79	90	88	56	61.5
27	108+ 92	97+ 92	110	109	-25	15
27C	115+105	112+109	72	96	2	27.5
28	150+148	162+143	81	84	29.5	73.5
28C	178+148	170+163	95	98	64.5	51.5

We have therefore:

28 pairs of I. Q's to be compared  
21 pairs of weights  
19 pairs of heights  
16 pairs of grip measurements  
20 pairs of tapping speeds  
24 pairs of Healy Completion scores.

Again, in only five pairs does the gain of the control exceed that of the test case, while in the remaining sixteen pairs the gains of the test cases are greater than those of their respective controls. The greatest loss of test as compared to control is 4.2 lbs., while the largest gain is 7.9. It would seem then, that after a six months' interval a child who has been operated on for adenoids and tonsils will tend to show a slightly greater increase in weight than a child who continues to suffer from the defects. The very small group renders this conclusion far from assured. Since it doubtless takes some little time to recover from the effects of the operation, and since there is comparatively little gain in weight in a six months' interval, it would be well to extend the experiment over another year. For the greater reliability of results, some degree of after-care should be given the operative cases, the control cases of course receiving the same treatment. While this was impracticable in the present study, it happened that three pairs of cases were members of a nutrition class, and therefore underwent some hygienic treatment. In one pair, (no. 11) the test case lost .3 of a pound, while the control gained .9. The test cases of pairs 7 and 10 gained .8 lb. and 2.1 lbs. respectively, over and above their controls. However, these three cases alone are of little significance.

A study of increase in height suffers even more than one of weight gains from the short interval which elapsed between measurements. Normally, there is very little growth in six months. There are only nineteen pairs of cases in this portion of the study, a fact which renders it of even less value. However, results are offered for what they are worth. The smallest increase in height (.3 in.) is in the test group, while the greatest growth (3.5 in.) is also in the test group. There is, however, a gain of 3.3 inches in the control group as well as one of only .4 inches. There are seven pairs in which the test group growth is less than that of the controls, one in which the two are equal, and in the remaining eleven the growth of the test cases exceeds that of the controls. The variability



TABLE III

Gain in weight, 6 months, 21 pairs

N*	Test Group (A)			Control Group (B)			Lbs.
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	A—B
8	60.8	63.3	2.5	45.4	52.1	6.7	—4.2
2	40.9	42.9	2.0	52.3	57.4	5.1	—3.1
28	71.3	75.5	4.2	74.2	79.8	5.6	—1.4
6	43.8	44.5	.7	50.6	52.5	1.9	—1.2
11	47.8	47.5	— .3	41.6	42.5	.9	—1.2
1	50.4	54.2	3.8	53.5	57.2	3.7	.1
7	39.9	41.0	1.1	38.4	38.7	.3	.8
9	50.6	53.2	2.6	59.8	61.4	1.6	1.0
12	48.0	52.5	4.5	41.0	44.5	3.5	1.0
14	56.0	62.0	6.0	81.9	86.0	4.9	1.1
5	45.0	47.0	2.0	56.2	57.0	.8	1.2
21	70.7	76.5	5.8	62.4	67.0	4.6	1.2
4	51.1	54.2	3.1	49.4	51.0	1.6	1.5
10	48.9	51.4	2.5	47.1	47.5	.4	2.1
16	56.0	57.0	1.0	51.2	50.0	—1.2	2.2
24	58.5	62.5	4.0	60.0	61.0	1.0	3.0
3	55.0	59.5	4.5	61.5	62.9	1.4	3.1
25	55.5	59.5	4.0	63.3	63.8	.5	3.5
19	90.0	96.3	6.3	52.4	54.4	2.0	4.3
13	90.0	98.0	8.0	74.7	76.8	2.1	5.9
26	63.8	74.5	10.7	64.2	67.0	2.8	7.9
Av.	56.86	60.61	3.76	56.24	58.60	2.39	1.37
M			3.8			1.9	1.2
75%ile			5.8			4.6	3.1
25%ile			2.0			.9	.1
Q			1.9			1.85	1.5
P. E. (distribution)			1.76			1.39	1.63
P. E. (average)			±.38			±.30	±.48

Av. = 2.85 P.E.

M. = 2.80 P.E.

of the test group growth is greater than that of the control group. The three nutrition pairs show the following records of growth,—in number 7, the test case shows a growth of .6 in. more than his control. Number 10 is the pair in which the growth is equal. In number 11 the test case again exceeds in growth by .6 of an inch.

More reliable than height and weight considered separately, as an index of physical welfare, is weight in relation to height and age. Table V shows the improvement in this relationship for the two groups. The numbers in columns 1, 2, 4 and 5 show the per cent under or over weight of the individual cases, in relation to their respective heights and ages. The authority upon which the figures are based, is the table published by the American Child Health Association, giving standard weights for height and age in boys.

\*Numbers refer to cases as listed on Table II.

There was an average loss of .28 per cent in the weight-height-age relationship for the test group, and of 2.11 per cent for the control group. The average improvement of the test group in excess of the control group is, then, 1.83 per cent. The median improvement of test group over and above control is 4.00 per cent. The test group is more variable than the control in improvement. The greatest improvement, 8 per cent, is found in both groups.

TABLE IV  
Gain in Height—6 Months, 19 Pairs

N*	Test Group (A)			Control Group (B)			Inches
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	A—B
9	48.1	49.4	1.3	51.9	55.2	3.3	—2.0
4	47.5	49.2	1.7	48.9	52.0	3.1	—1.4
5	44.9	45.2	.3	46.6	48.1	1.5	—1.2
6	43.1	43.9	.8	45.4	47.3	1.9	—1.1
2	42.5	44.1	1.6	45.2	47.0	1.8	— .2
28	54.9	55.8	.9	53.4	54.5	1.1	— .2
16	51.6	52.3	.7	48.7	49.5	.8	— .1
10	46.1	47.7	1.6	45.6	47.2	1.6	0
25	50.0	50.8	.8	50.2	50.9	.7	.1
1	46.0	47.6	1.6	46.4	47.8	1.4	.2
3	47.0	48.4	1.4	51.7	52.9	1.2	.2
8	50.8	51.8	1.0	46.8	47.6	.8	.2
14	51.0	51.6	.6	57.9	58.3	.4	.2
7	42.9	44.8	1.9	41.9	43.2	1.3	.6
11	45.8	47.7	1.9	43.6	44.9	1.3	.6
19	57.7	59.0	1.3	46.7	47.2	.5	.8
26	51.6	54.3	2.7	51.4	52.3	.9	1.8
22	53.0	56.4	3.4	56.8	58.1	1.3	2.1
13	61.5	65.0	3.5	56.8	57.8	1.0	2.5
Av.	49.26	50.79	1.53	49.20	50.62	1.36	.16
M			1.4			1.3	.2
75%ile			1.90			1.75	.65
25%ile			.78			.8	— .43
Q			.56			.48	.54
P. E. (distribution)			.53			.44	.44
P. E. (average)			±.12			±.10	±.16
						Av.=1	P. E.
						M =1.25	P. E.

The greatest loss, 10 per cent, is in the control group. Eight cases show a loss in comparison to their controls, and nine reveal a gain. On the whole, there is some significance in the small net improvement manifested by the test group. The average is 2.02 P. E.'s, and the median 4.40 P. E.'s.

The dynamometer results show no gain in strength of grip six months after operation. Indeed the average of the gains

\*Numbers refer to cases as listed on Table II.

of the operative cases is slightly less than the average gain of the controls. Comparing the test group with the control, we find the average of the differences to be  $-.24$ . But the variability is so high ( $P. E. = +.48$ ) as to render this figure unreliable. The greatest loss in strength of grip is found in the control group, but the greatest gain is also in this group. Seven cases in the test group show a loss, as compared with only three control cases. In eight, or one-half of the sixteen cases, the control member of a pair gained more than the test member. Considering the three pairs of nutrition cases, we find that in pair number 7 the test case loses 1.5 Kg. when compared with the control; and in pair number 10, 6.5 Kg., while the test case in pair 11 gains 4 Kg. The conclusion from the data would seem to be that, within the space of six months at any rate, operation for adenoids and tonsils brings about no increase in strength of grip.

TABLE V

Showing change in per cent over or underweight for height and age,  
18 pairs

	A				B		A-B
N*	1	2	3	4	5	6	7
8	- 1	- 1	0	-13	- 5	+ 8	- 8
11	- 3	-12	-9	- 9	-11	- 2	- 7
10	- 1	- 5	-4	- 6	- 8	- 2	- 2
28	- 7	- 7	0	0	+ 2	+ 2	- 2
13	-13	-19	-6	-11	-15	- 4	- 2
19	+ 6	+ 8	+2	+ 1	+ 5	+ 4	- 2
6	+ 2	- 3	-5	+ 5	+ 1	- 4	- 1
14	-13	- 9	+4	- 5	0	+ 5	- 1
7	- 8	-15	-7	- 7	-14	- 7	0
2	- 8	- 6	+2	+ 9	+10	+ 1	+ 1
3	+ 6	+ 8	+2	- 5	- 8	- 3	+ 5
16	-16	-15	+1	-13	-17	- 4	+ 5
25	- 8	- 7	+1	+ 4	- 1	- 5	+ 6
5	- 4	- 1	+3	+ 8	+ 4	- 4	+ 7
9	- 8	-11	-3	- 7	-17	-10	+ 7
26	- 1	+ 5	+6	+ 1	0	- 1	+ 7
4	- 7	- 7	0	-15	-23	- 8	+ 8
1	0	+ 8	+8	+ 8	+ 4	- 4	+12
Av.	- 4.67	- 4.94	-.28	- 3.06	- 5.17	- 2.11	+ 1.83
M			+ .5			- 3.5	+ 4.00
75%ile			+2			0	6.5
25%ile			-2			- 4.5	- 2
Q			2			2.25	4.25
P. E. (distribution)			3			2.39	1.33
P. E. (average)			±.71			± 57	± 91
						Av.=2.02	P. E.
						M.=4.40	P. E.

\*Numbers refer to cases as listed on Table II.



Is there, after operation, an improvement in motor control and attention, and a lessening of fatiguability as these may be demonstrated in the tapping test? Table VI gives the number of taps in the first half minute of tapping for both groups before and after the six months interval. The test group suffers an average loss of 2.24 taps, and a median loss of 2. The average loss of the control group is 2.33, and the median 2.

TABLE VI  
Gain in Grip—6 Months—16 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
10	12.5	11	—1.5	10	15	.5	—6.5
27	9	6	—3	8	9	1	—4
1	13	12	—1	11	13	2	—3
14	16	15	—1	22	24	2	—3
2	9	9	0	10	12	2	—2
9	10.5	13	2.5	16.5	21	4.5	—2
7	7	6.5	— .5	9	10	1	—1.5
16	19	17.5	—1.5	10	10	0	—1.5
3	12.8	14.5	1.7	14	15	1	.7
13	26.5	28.5	2	22	23	1	1
24	19	20	1	15	15	0	1
19	22	21	—1	15	12	—3	2
22	18	22.5	4.5	19.5	21.5	2	2.5
11	11	15	4	11.5	11.5	0	4
21	16.5	16.5	0	19	15	—4	4
23	11.5	15.5	4	14.5	14	— .5	4.5
Av.	14.58	15.22	.62	14.19	15.06	.875	— .24
M			0			1	—1.0
75%ile			3			2	2.25
25%ile			—1			0	—2.5
Q			2			1	2.38
P. E. (distribution)			1.58			1.02	2.49
P. E. (average)			±.40			±.26	±.48
							Av.— .50 P. E.
							M.—2.08 P. E.

There is practically no change then in the tapping ability of either group. The high unreliability of the difference (P. E. =  $\pm 3.10$ ) is noteworthy. It would seem that incidental causes have a much greater effect upon tapping ability than can be demonstrated as resulting from the removal of adenoids and tonsils.

Use of the tapping test as a measure of the decrease in tendency to fatigue similarly brings out no indication of any improvement in the operative group of cases. The measure of

\*Numbers refer to cases as listed on Table II.

fatigue was taken as a ratio; namely, the number of taps in the first, minus the number in the second half minute over the number of taps in the first half minute. Then, if there is a greater number of taps in the second, the ratio will be minus, indicating that fatigue effect is so small as to be overcome by practice effect. This was a fact in only four cases. Since what we are measuring is improvement, the ratio for test 2 is subtracted from the ratio for test 1 to find the gain in overcoming fatigue. Table VIII shows the average gain for group one to be  $-.0196$ , and the median  $-.045$ . That is, there is an average increase in fatiguability of  $.0196$  units and a median increase of  $.045$  with a P. E. of  $\pm .02$ . This increase in fatiguability occurs also in the control group, average 0, and Median  $.03$  with P. E. of  $\pm .03$ . The average gain of test group over control group is  $-.02$  and the median gain is  $-.015$ . Again variability is relatively large, P. E. being  $1.04$ , so that the median and average gains are  $-.50$  P. E. and  $-.38$  P. E. respectively.

We may say, then, that the capacities brought out by the tapping test seem to undergo no improvement in six months after removal of adenoids and tonsils.

The main line of interest in the present experiment lay with the relation of adenoid and tonsil defects to general intelligence. The results of the two tests dealing more specifically with this side of the problem are here set forth. Table IX shows the I. Q's. of the two groups before and after the six months' interval, together with changes plus or minus in I. Q., and a comparison of the separate pairs in respect to improvement.

We find that the test group shows an average gain in I. Q. of 2.25 points. The median gain is 2 points, the total range 18 points and P. E. of the average is  $\pm .99$ . The control group shows an average gain very slightly higher, 3.25 points, the median gain being 3. The range in this case is 32 points, but P. E. is only  $\pm .47$ . The average of the compared gains of separate pairs is  $-1.035$ . These numbers are so small as to be insignificant. Actually, we may say that the operative group as a whole showed no gain over the control group. If we examine individual cases we find that the greatest loss in I. Q. was in the control group, (8 points) but the greatest gain (24 points) also appears in this group. In the test group 11 cases

TABLE VII

Gain in number of taps in one-half minute, 21 pairs—right hand

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
14	190	138	—52	175	175	0	—52
10	68	88	20	70	135	65	—45
23	150	141	—9	122	140	18	—27
1	135	120	—15	106	115	9	—24
9	149	135	—14	144	150	6	—20
21	152	132	—20	140	138	—2	—18
22	164	183	19	120	157	37	—18
24	157	142	—15	155	155	0	—15
27	108	97	—11	115	112	—3	—8
20	105	110	5	150	155	5	0
15	172	170	—2	140	137	—3	1
26	137	138	1	125	125	0	1
25	140	150	10	148	151	3	7
3	136	139	3	135	129	—6	9
13	160	142	—18	150	122	—28	10
18	133	135	2	100	84	—16	18
28	150	162	12	178	170	—8	20
8	113	110	—3	131	101	—30	27
11	125	98	—27	155	101	—54	27
2	105	112	7	152	114	—38	45
17	90	150	60	125	121	—4	64
Av.	135.19	132.95	—2.24	136.47	134.14	—2.33	.09
M			—2			—2	0
75%ile			6.5			4.5	16.0
25%ile			—15			—14	19.5
Q			10.75			9.25	17.75
P. E. (distribution)			12.24			7.33	18.09
P. E. (average)			±2.66			±1.59	±3.10
						Av.=	.03 P. E.
						M.=	.0 P. E.

lost in I. Q., as compared with 7 in the control group. Thirteen test cases lost in comparison with their respective controls. Two gained equally with their controls, and the remaining thirteen showed a larger gain. In regard to the three pairs taken from the nutrition class, number 7 gained 8 points and his control, 4. Number 10 lost a point and his control lost 4, while number 11 lost 3 points with a gain of 1 point by his control. So that these cases, in spite of most favorable conditions, show no consistent gain in I. Q.

The results of the Healy tests are similar. There is a slightly higher average gain in the control group. The test group contains eight cases which made a poorer score at the end of the interval, the control group six. The range of gains is from —22 to +44, or 66 points, in the test group, while in the control group the gains range from —14 to +41.5 or 55.5

\*Numbers refer to cases as listed on Table II.



TABLE VIII

Decrease in fatigue in tapping—Difference in rates of second half minute over first half minute. Sixteen pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
13	— .03	.05	— .08	.27	— .28	.55	— .63
28	.01	.11	— .10	.17	.04	.13	— .23
20	.10	.15	— .05	.20	.04	.16	— .21
10	— .09	.07	— .16	.23	.19	.04	— .20
17	.01	.33	— .32	.07	.20	— .13	— .19
25	.09	.21	— .12	.09	.11	— .02	— .10
11	.28	.11	.17	.19	— .06	.25	— .08
15	.03	.08	— .05	.18	.16	.02	— .07
22	.10	.23	— .13	.03	.19	— .16	.03
27	.15	.05	.10	.09	.03	.06	.04
14	.09	.06	.03	.02	.06	— .04	.07
18	.14	.18	— .04	.01	.12	— .11	.07
24	.13	.11	.02	.13	.35	— .22	.24
26	.18	.15	.03	.16	.37	— .21	.24
23	.21	.04	.17	.06	.21	— .15	.32
21	.27	.05	.22	.03	.20	— .17	.39
Av.	.104	.124	— .020	.121	.121	.0	— .020
M			— .045			— .03	— .015
75%ile			.03			+ .04	.07
25%ile			— .12			— .16	— .20
Q			.075			.10	.135
P. E. (distribution)			.09			.11	.05
P. E. (average)			± .02			± .03	± .04
						Av. = — .50 P. E.	
						M. = — .38 P. E.	

points. Seventeen of the operative cases showed a smaller gain than their respective controls. The three pairs of cases from the nutrition class show the following gains:—pair 7; the test case loses 22 points, the control gains 23 points; pair 10, test case gains 18.5, but control gains 38 points; pair 11, test case gains 14 points, and control gains 25.5 points. From this test then, we can find no general tendency for cases operated on to improve in intelligence in excess of improvement in a control group which was not so treated.

This question presents itself:—is there any relationship between improvement in physical well-being as revealed in weight, and improvement in intelligence? If, as has been supposed, adenoids and diseased tonsils cause mental retardation indirectly through physical deprivation, it would seem as though greater improvement in intelligence after operation should accompany greater improvement in weight, and smaller intelligence gain should accompany slighter gain in weight.

\*Numbers refer to cases as listed on Table II.

In order to determine whether this was true for our cases, improvement in I. Q. was correlated with gain in weight, for the test group. The order of merit method was used, and the formula  $\rho = 1 - \frac{6 \sum D_o}{n(n^2-1)}$  where  $f = 2 \sin \left( \frac{\pi}{6} \rho \right)$ . The resulting value of  $r$  was  $-.10$  with unreliability of  $.226$ , calculated by the formula  $\sigma_{t.r} - \text{obt.}r = \frac{1.05(1-r^2)}{\sqrt{n}}$ . There is therefore no correlation between improvement in intelligence and gain in weight.

TABLE IX

Improvement in I. Q., 28 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
27	110	109	— 1	72	96	24	—25
20	106	102	— 4	118	131	13	—17
8	91	86	— 5	98	104	6	—11
24	131	124	— 7	89	92	3	—10
21	64	67	3	86	97	11	— 8
3	94	91	— 3	82	85	3	— 6
12	98	95	— 3	98	101	3	— 6
23	85	94	9	81	96	15	— 6
11	103	100	— 3	101	102	1	— 4
15	97	94	— 3	78	79	1	— 4
18	98	98	0	90	94	4	— 4
4	96	96	0	83	85	2	— 2
26	80	76	— 4	90	88	— 2	— 2
16	65	73	8	74	82	8	0
28	81	84	3	95	98	3	0
5	95	99	4	114	117	3	1
2	107	114	7	91	96	5	2
10	110	109	— 1	104	100	— 4	3
17	71	77	6	96	99	3	3
7	91	99	8	95	99	4	4
1	82	83	1	80	76	— 4	5
6	95	101	6	88	89	1	5
19	96	101	5	98	98	0	5
9	83	93	10	87	90	3	7
25	77	76	— 1	145	137	— 8	7
13	70	78	8	66	64	— 2	10
22	91	100	9	63	62	— 1	10
14	96	107	11	140	137	— 3	14
Av.	91.53	93.78	2.25	92.93	96.21	3.285	—1.035
M			2			3	—1
75%ile			7			4	5
25%ile			—3			—1	—6
Q			5			2.5	5.5
P. E. (distribution)			5.25			2.5	5
P. E. (average)			±.99			±.47	±1.10
						Av.—.94	P. E.
						M. =—.99	P. E.

\*Numbers refer to cases as listed on Table II.

TABLE X

Improvement in Performance of Healy Test, 24 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
7	— 6	—28	—22	4.5	27.5	23	—45
13	43	42	— 1	— 1.5	30.5	—32	—33
17	29.5	12	—17.5	1.5	15	13.5	—31
15	7	25	18	1	42.5	41.5	—23.5
26	22.5	7	—15.5	56	61.5	5.5	—21.5
10	—12	6.5	18.5	27	65	38	—19.5
19	57.5	49	— 8.5	—22	—11	11	—19.5
24	54.5	63	8.5	31.5	59.5	28	—19.5
20	0	—11	—11	30	35	5	—16
9	3.5	10.5	7	34	55	21	—14
11	— 8	6	14	—29	— 3.5	25.5	—11.5
16	49	47.5	— 1.5	30	37	7	— 8.5
1	—25	2	27	—50	—16	34	— 7
22	48.5	43.5	— 5	34.5	33.5	— 1	— 4
18	—13.5	—12	1.5	—32	—28	4	— 2.5
23	49.5	68	18.5	4	25	21	— 2.5
3	21.5	22.5	1	17	19	2	— 1
12	20	21	1	—10	—12	— 2	3
14	12.5	48.5	36	— 5	25.5	30.5	6.5
27	—25	15	40	2	27.5	25.5	14.5
2	28.5	30	1.5	3	—11	—14	15.5
25	8	25	17	29.5	29	— .5	17.5
21	20	32	12	70.5	58.5	—12	24
28	29.5	73.5	44	64.5	51.5	—13	57
Av.	17.29	24.94	7.64	12.12	25.69	13.56	— 5.85
M.			4.25			12.25	— 7.75
75%ile			18			25.5	3
25%ile			— 5			— .5	—19.5
Q			11.5			13	11.25
P. E. (distribution)			10.6			10.56	13.65
P. E. (average)			±2.16			±2.16	±3.05
						Av.—	—1.92 P. E.
						M.—	—2.54 P. E.

Similarly, it might be thought that the children who had suffered from the defects for a comparatively short time, might reveal greater improvement in intelligence after six months than those who had been afflicted for a longer space of time. We had no way of knowing definitely how long the defects had been present in the cases studied. Roughly, though, we may say that in general the older boys have had defective tonsils and adenoids for a longer time than the younger ones, and that the older the boy, the older the defect. On this basis, if correlation of youth with gain in I. Q. should give a larger positive value for  $r$ , we might be justified in saying that the younger boys, who have been handicapped for a lesser period, show greater mental recuperation than their

\*Numbers refer to cases as listed on Table II.



older companions. Such a correlation was attempted in the test group, correlating age at the first test with gain in I. Q. The same methods and formulae were used as in the weight and intelligence comparison, the greatest gain in I. Q. being given first position, and the lowest age. The resulting value for  $r$  was  $-.24$ , with an unreliability of  $.186$ . The relationship would appear to be in the other direction but it is so small, with an unreliability measure so large as to be insignificant. Once more, then, we find in our results no correspondence between recency of defect and quick mental recovery.

TABLE XI

Showing percentile ratings of the members of the two groups at the beginning and end of the six months' interval

	Weight		Height		Grip		Tapping	
1	.29	.44	.25	.40	.47	.40	.51	.33
1C	.43	.54	.27	.45	.33	.47	.17	.30
2	.04	.10	.04	.11	.16	.16	.16	.25
2C	.38	.55	.20	.32	.25	.40	.80	.28
3	.46	.59	.32	.49	.44	.56	.52	.58
3C	.65	.69	.74	.81	.54	.67	.51	.42
4	.33	.44	.39	.52	.16	—	.13	—
4C	.27	.32	.51	.78	.18	—	.19	—
5	.16	.20	.18	.20	.33	—	.23	—
5C	.50	.53	.28	.47	.40	—	.87	—
6	.12	.15	.06	.10	.10	—	.23	—
6C	.31	.40	.22	.38	.27	—	.41	—
7	.03	.07	.05	.15	.07	.06	.40	.27
7C	.01	.02	.03	.07	.16	.24	.16	—
8	.62	.71	.63	.75	.24	—	.27	.23
8C	.18	.37	.31	.40	.67	.72	.43	.11
9	.31	.41	.47	.53	.26	.47	.71	.51
9C	.60	.65	.76	.89	.76	.89	.63	.78
10	.26	.35	.26	.43	.44	.33	.01	.04
10C	.21	.22	.23	.37	.24	.67	.02	.51
11	.24	.28	.25	.43	.33	.67	.40	.08
11C	.08	.09	.09	.17	.36	.36	.86	.11
12	.25	.40	.15	—	.53	—	.08	—
12C	.06	.15	.02	.08	.04	.01	.12	.03
13	.95	1.00	.99	1.00	.99	1.00	.90	.66
13C	.87	.89	.93	.96	.94	.96	.78	.36

TABLE XI (Continued)

Showing percentile ratings of the members of the two groups at the beginning and end of the six months' interval

	Weight		Height		Grip		Tapping	
14	.49	.66	.67	.74	.72	.67	1.00	.57
14C	.92	.93	.96	.98	.93	.98	.97	.97
15	.56	—	.68	—	.70	—	.95	.94
15C	.79	.80	.59	.66	.67	.70	.62	.54
16	.49	.53	.74	.79	.81	.77	.69	—
16C	.34	.28	.50	.54	.24	.24	.69	.51
17	.53	—	.50	—	.53	—	.05	.78
17C	.17	—	.15	—	.24	.11	.40	.34
18	.57	—	.38	—	.33	—	.45	.51
18C	.17	.20	.30	.34	.10	.06	.09	.03
19	.95	.99	.94	.98	.94	.88	.93	—
19C	.38	.45	.30	.37	.67	.40	.09	.31
20	.13	—	.37	—	.33	—	.16	.23
20C	.63	.75	.56	.88	.67	.44	.78	.86
21	.83	.89	.84	—	.76	.76	.80	.44
21C	.67	.77	.55	.61	.81	.67	.62	.57
22	.85	—	.81	.91	.78	.95	.92	.99
22C	.83	.92	.93	.97	.82	.91	.33	.89
23	.36	—	.45	—	.36	.70	.78	.63
23C	.73	.74	.70	.82	.55	.53	.36	.62
24	.59	.69	.67	—	.81	.85	.89	.66
24C	.62	.63	.59	.71	.67	.67	.86	.86
25	.47	.60	.57	.63	.53	—	.62	.78
25C	.73	.74	.60	.64	.44	.91	.70	.79
26	.74	.89	.74	.85	.53	—	.54	.57
26C	.76	.81	.70	.79	.84	.76	.40	.40
27	.11	—	.23	—	.16	.04	.18	.06
27C	.07	.12	.12	.23	.10	.16	.30	.25
28	.86	.90	.87	.90	.97	.88	.78	.91
28C	.87	.93	.83	.86	.88	—	.98	.94

TABLE XI (Continued)

Showing percentile ratings of the members of the two groups at the beginning and end of the six months' interval

N*	I. Q.		Healy		Total	Possible Gain	Average
1	.25	.27	.10	.29	.30	415	.05
1C	.21	.15	.01	.12	.53	462	.089
2	.84	.89	.60	.66	.38	416	.063
2C	.45	.51	.29	.18	— .07	358	— .011
3	.49	.44	.49	.51	.40	328	.066
3C	.25	.32	.44	.45	.30	287	.05
4	.59	.59	.38	—	.22	168	.073
4C	.27	.32	.71	—	.42	195	.14
5	.53	.71	.10	—	.42	213	.14
5C	.89	.90	.77	—	.24	133	.08
6	.53	.77	.02	—	.55	229	.183
6C	.37	.38	.04	—	.27	210	.09
7	.45	.71	.22	.07	.37	478	.062
7C	.53	.71	.34	.59	.74	393	.148
8	.45	.34	.70	—	— .50	303	— .125
8C	.67	.81	.33	.52	.10	241	.025
9	.27	.46	.23	.39	.71	375	.118
9C	.35	.40	.73	.89	.72	217	.12
10	.88	.86	.15	.35	.34	400	.056
10C	.81	.73	.57	.97	1.39	392	.218
11	.79	.73	.20	.34	.21	379	.035
11C	.77	.79	.05	.23	— .44	379	— .073
12	.67	.53	.47	.49	— .11	161	— .036
12C	.67	.77	.19	.15	.19	208	.032
13	.09	.19	.79	.78	.02	129	.003
13C	.06	.05	.24	.67	.06	178	.01
14	.59	.84	.41	.84	.69	212	.115
14C	.97	.96	.21	.56	.41	104	.068
15	.62	.49	.38	.55	.12	106	.04
15C	.19	.20	.26	.79	.66	193	.11
16	.05	.12	.86	.82	.15	205	.03
16C	.13	.25	.66	.76	.24	313	.04
17	.09	.17	.63	.40	.66	223	.220
17C	.59	.71	.27	.43	.21	174	.053
18	.67	.67	.13	.15	.08	175	.027
18C	.40	.49	.04	.07	.18	247	.03

\*Numbers refer to cases as listed on Table II.



TABLE XI (Continued)

Showing percentile ratings of the members of the two groups at the beginning and end of the six months' interval

N*	I. Q.		Healy		Total	Possible Gain	Average
19	.59	.69	.91	.85	.16	67	.032
19C	.67	.67	.11	.18	.16	287	.027
20	.82	.79	.25	.18	— .06	177	— .02
20C	.91	.94	.66	.75	.20	65	.066
21	.05	.07	.47	.69	— .16	209	— .032
21C	.34	.62	.99	.92	.46	157	.077
22	.45	.73	.84	.81	.87	120	.174
22C	.03	.02	.74	.72	.74	275	.123
23	.32	.49	.86	.98	.65	168	.163
23C	.23	.59	.33	.50	1.26	253	.21
24	.94	.92	.88	.95	— .06	89	— .012
24C	.38	.46	.68	.93	.54	179	.09
25	.17	.15	.37	.55	.49	280	.098
25C	1.00	.97	.63	.60	.53	134	.088
26	.21	.15	.51	.36	0	226	0
26C	.40	.37	.90	.94	.04	184	.007
27	.88	.86	.10	.43	.05	268	.013
27C	.11	.59	.29	.59	1.49	320	.25
28	.23	.29	.63	1.00	.59	153	.097
28C	.53	.67	.96	.87	.24	83	.048

Table XI expresses the results of Table II, with the scores given in percentile values. In each test, the group was taken as composed of the two scores of every individual—the total number of scores in tests and retests, eliminating those scores where the other member of the pair was lacking, or where no retest was given. Thus case number 1 was just within the lowest 27% of the group in weight at the first weighing, but had advanced to the 44 percentile at the second. In height he gained from the 25 percentile to the 40 percentile. His total gain in all tests is 30 percentile out of a possible 415, and the average gain is .05. The reader may see by scanning the table that the gains in the test group are practically equalled by those in the control group. There seems to be no consistent relationship between a low score in the first test and a large gain. This is true even though the method of calculation tends to minimize gains at the high end of the group, and losses at the

low end. In table XII this may be seen more clearly in respect to I. Q. and the results for all the tests taken together with the I. Q. weighted by being counted twice. A large possible gain indicates that the score at the first testing was low, and vice versa. Considering I. Q. values, the largest possible gain in the test group was 95 per cent of the group. This occurred twice, in one case the actual gain being 7% of the group and in the other 2%. In the control group, the largest possible gain was 97% of the group, but actually this case fell 1% of the group. If we correlate possible gain with actual gain for each group, using the formula  $r = 2\sin\left(\frac{\pi}{6}\rho\right)$  when  $\rho = 1 - \frac{6\sum D^2}{n(n^2-1)}$  we get a coefficient of correlation .36 in the test group, and .19

TABLE XII

Showing gains in percentile rating for I. Q., and for a total of all the tests with I. Q. weighted by being counted twice.

	I. Q.		A possible gain	B actual gain	Total		Av. Gain
	1st P.R.	2d P.R.			possible gain	actual gain	
1	25	27	75	2	415	30	5
1C	21	15	79	— 6	462	53	8.9
2	84	89	16	5	416	38	6.3
2C	45	51	55	6	358	— 7	— 1.1
3	49	44	51	— 5	328	40	6.6
3C	25	32	75	7	287	30	5
4	59	59	41	0	168	22	7.3
4C	27	32	73	5	195	42	14
5	53	71	47	18	213	42	14
5C	89	90	11	1	133	24	8
6	53	77	47	24	229	55	18.3
6C	37	38	63	1	210	27	9
7	45	71	55	26	478	37	6.2
7C	53	71	47	18	393	74	14.8
8	45	34	55	—11	203	—50	—12.5
8C	67	81	33	14	241	10	2.5
9	27	46	73	19	375	71	11.8
9C	35	40	65	5	217	72	12
10	88	86	12	— 2	400	34	5.6
10C	81	73	19	— 8	392	139	21.8
11	79	73	21	— 6	379	21	3.5
11C	77	79	23	2	379	—44	— 7.3

TABLE XII (Continued)

Showing gains in percentile rating for I. Q., and for a total of all the tests with I. Q. weighted by being counted twice.

I. Q.					Total		
	1st P.R.	2d P.R.	A possible gain	B actual gain	possible gain	actual gain	Av. Gain
12	67	53	33	— 6	161	—11	— 3.6
12C	67	77	33	10	208	19	3.2
13	9	19	91	10	129	2	.3
13C	6	5	94	— 1	178	6	1
14	59	84	41	25	212	69	11.5
14C	97	96	3	— 1	104	41	6.8
15	62	49	38	—13	106	12	4
15C	19	20	81	1	193	66	11
16	5	12	95	7	205	15	3
16C	13	25	87	12	213	24	4
17	9	17	81	8	223	66	22
17C	59	71	41	12	174	21	5.3
18	67	67	33	0	175	8	2.7
18C	40	49	60	9	247	18	3
19	59	69	41	10	67	16	3.2
19C	67	67	33	0	287	16	2.7
20	82	79	18	— 3	177	— 6	— 2
20C	91	94	9	3	65	20	6.6
21	5	7	95	2	209	—16	— 3.2
21C	34	62	66	28	157	46	7.7
22	45	73	55	28	120	87	17.4
22C	3	2	97	— 1	275	74	12.3
23	32	49	68	17	168	65	16.3
23C	23	59	77	36	253	126	21
24	94	92	6	— 2	89	— 6	— 1.2
24C	38	46	62	8	179	54	9
25	17	15	83	— 2	280	49	9.8
25C	100	97	0	— 3	134	53	8.8
26	21	15	79	— 6	226	0	0
26C	40	37	60	— 3	184	4	.7
27	88	86	12	— 2	268	5	1.3
27C	11	59	89	48	320	148	25
28	23	29	77	6	153	59	9.7
28C	53	67	47	14	83	24	4.8



in the control group. With the small number of cases involved the probable error is too great to allow either of these measures as indicative of relationship. We may say, then, that there is no definite tendency for those of low I. Q. to improve in six months after operation to a greater degree than those of higher I. Q.

Finally, in order to compare the results of the various tests, the measures of the gains of the test group in excess of the control were, for each test, expressed in terms of P. E. The averages and medians of these measures are collected in Table XIII. They show a very slight tendency toward gain in weight, height, and weight-height-age relationship; neither improvement nor loss in grip, tapping fatigueability and I. Q., and a rather curious tendency to loss in the Healy scores. This latter is very probably not a true measure since performance in the Healy Picture Completion test shows a rather high variability, and the cases are so few as to make the influence of single very high or low scores unduly great.

TABLE XIII

Showing improvement in various tests of operative group over and above such improvement in control group. Expressed in Terms of P.E.

	Weight	Height	Height- Weight	Grip	Tapping	Tapping fatigue	I. Q.	Healy
	P. E.	P. E.	P. E.	P. E.	P. E.	P. E.	P. E.	P. E.
Average	2.85	1.00	2.02	— .50	.03	— .50	— .94	— 1.92
Median	2.80	1.25	.55	— .83	.32	— .50	0	— 2.54

## CHAPTER IV

### MEASUREMENT OF IMPROVEMENT AFTER A SECOND INTERVAL OF SIX MONTHS

In view of the fact that one of the experimenters\* found improvement in school work when her study was extended to cover a second time interval after operation, it was deemed advisable to similarly extend the present investigation in order to determine whether our operated cases showed any improvement after twelve months. To this end, the fifty-six children composing the final groups of the first study, were sought after a second interval of about six months. Conditions made it impossible to give all the retests exactly twelve months from the time of the operation. As a matter of fact, the period ranges from ten to seventeen months. An effort was made to keep the interval between tests equal for the two members of a given pair.

The same tests were given as in the first study. About half of the testing was done by one of the former examiners, but she was obliged to turn the work over to another before it had been completed. The second examiner was highly recommended, and had had training and practical experience in the giving of tests. She was instructed in the methods which had been employed previously, so that conditions were as far as possible kept constant.

The results of the tests are collected in Table XIV. In the first column is given the length of the time interval for each case. It may be seen that the final group was composed of forty-two children, forming twenty-one pairs. There were fifteen pairs which received a second rating in weight; thirteen in height; thirteen in grip; fifteen in tapping, eleven in fatigue as shown by tapping, twenty-one in I. Q., and eighteen in the Healy Test. These numbers while they are smaller than we could wish, would seem to be great enough to indi-

---

\*A. H. Mac Phail, *Adenoids and Tonsils: A study showing how the Removal of Enlarged or Diseased Tonsils Affects a Child's Work in School*. Ped. Sem., June, 1920, pp. 188-194.

TABLE XIV

Results of the Tests after an Interval of from 10 to 17 Months

N*	Mos.	Weight		Height		Grip	
		Test 1	Test 3	Test 1	Test 3	Test 1	Test 3
1	15	50.4	63.5	46	49.7	13	18
1C	15	53.5	62.8	46.4	49.7	11	15
2	15	40.9	47.9	42.6	45.9	9	10
2C	17	52.3	65.5	45.2	49.8	10	15
3	16	55	67.5	47	50.2	12.8	13
3C	14	61.5	57.8	51.7	54	14	14.5
4	13	51.1	60.2	47.5	50.5	9	—
4C	13	49.4	54.2	48.9	50.8	9.5	—
7	14	39.9	45.8	42.9	45.6	7	6
7C	12	38.4	42.1	41.9	43.7	9	14
8	11	60.8	69.2	50.8	52.3	10	—
8C	11	45.4	57.9	36.8	48.7	15	—
10	11	48.9	56.7	46.1	48.6	12.5	12
10C	11	47.1	51.8	45.6	48.1	10	12
11	12	47.8	55	45.8	49.5	11	11
11C	12	41.6	47	43.6	46.8	11.5	7.5
12	12	48	66.5	44.8	—	14	—
12C	11	41	69.6	41.5	—	6	—
13	12	90	112	61.3	61.8	26.5	28.5
13C	12	74.7	88	56.8	60.0	22	27
14	12	56	66	51	53.3	16	17
14C	12	81.9	98	57.9	59.5	22	22.5
15	12	57.5	—	51.1	—	15.5	—
15C	10	67.2	—	50.1	—	15	—
16	12	56	60.3	51.6	53.5	19	18.5
16C	11	51.2	55	48.7	50.1	10	10
18	12	58.2	—	47.3	—	18	—
18C	11	45.3	—	46.7	—	8	—
19	12	90	108	57.7	60.5	22	20
19C	11	52.4	59	46.7	48	15	18
20	12	44.2	—	47.2	—	11	—
20C	11	61.3	—	49.6	—	15	—
21	12	70.7	85.5	54.1	—	16.5	15
21C	10	62.4	69.2	49.6	—	19	17
23	12	51.7	—	47.8	—	11.5	15
23C	11	64.1	—	51.4	—	14.5	16
27	12	43.7	—	45.4	—	9	6
27C	11	41	—	44.6	—	8	9
28	12	71.3	78.5	54.9	56.3	23.5	—
28C	11	74.2	85.8	53.4	55.9	21	—

\*Numbers refer to cases as listed on Table II.



TABLE XIV (Continued)

N*	Tapping		I. Q.		Healy	
	Test 1	Test 3	(1)	(3)	(1)	(3)
1	135	142	82	93	—25	11
1C	106	134	80	85	—50	11
2	105	135	107	113	28.5	24.5
2C	152	139	91	86	3	19.5
3	136	144	94	91	21.5	15.5
3C	135	135	82	96	17	25.5
4	103	—	96	111	8.5	—
4C	109	—	83	102	33	—
7	125	—	91	93	— 6	16
7C	105	—	95	112	4.5	11
8	113	128	91	92	32.5	—
8C	131	121	98	111	4	—
10	68+ 74	145+106	110	116	—12	11
10C	70+ 74	148+124	104	107	27	48.5
11	125+ 90	120+125	103	102	— 8	15.5
11C	155+125	102+112	101	95	—29	—20
12	98+ 69	—	98	86	20	1
12C	102	—	98	90	—10	41.5
13	160+165	176+187	70	61	43	62.5
13C	150+109	188+174	66	60	— 1.5	21.5
14	190+172	228+215	96	102	12.5	77
14C	175+152	165+186	140	138	— 5	48.5
15	172+167	192+186	97	97	7	19
15C	140+115	145+133	78	98	1	54.5
16	145+131	—	65	74	49	79
16C	145+ 99	—	74	81	30	45.5
18	133+115	126+145	98	101	—13.5	13.5
18C	100+ 99	108+ 92	90	92	—32	—35
19	168+136	—	96	97	57.5	60.5
19C	100+115	—	98	90	—22	—15
20	105+115	122+118	106	116	0	55
20C	150+120	154+154	118	140	30	48.5
21	152+111	154+155	64	66	20	38
21C	140+136	174+150	86	93	70.5	88
23	150+119	157+157	85	80	49.5	62.5
23C	122+115	141+141	81	88	4	64
27	108+ 92	114+ 95	110	112	—25	25.5
27C	115+105	101+118	72	98	2	39.5
28	150+148	176+168	81	83	29.5	77.5
28C	178+148	172+157	95	94	64.5	83.5

cate any very consistent tendency toward improvement. The question, whether or not the results are affected by the differences in time interval, will be considered later.

In weight, the test group showed an average gain of 11.013 pounds, with a median of 9.1 (Table XV). The average gain of the control group was 9.113 pounds and the median 6.8. The gains in the test group are less variable than those of the control. The average of the gains of the test group in excess of those of the control is 1.9 pounds, and the median is 2.2 pounds; while the unreliability of the difference is  $\pm 1.46$ . The average, then, is only 1.30 P. E. and the median 1.51 P. E.

If we turn to Table III and compare the results there set forth with the results at the end of the second period, we find the gains of the test group exceed those of the control in the following manner.

TABLE XV  
Weight, Second Retests, 15 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
12	48	66.5	18.5	41	69.6	28.6	—10.1
2	40.9	47.9	7	52.3	65.5	13.2	— 6.2
14	56	66	10	81.9	98	16.1	— 6.1
28	71.3	78.5	7.2	74.2	85.8	11.6	— 4.4
8	60.8	69.2	8.4	45.4	57.9	12.5	— 4.1
16	56	60.3	3.7	51.2	55	3.8	— .1
11	47.8	55	7.2	41.6	47	5.4	1.8
7	39.9	45.8	5.9	38.4	42.1	3.7	2.2
10	48.9	56.7	7.8	47.1	51.8	4.7	3.1
1	50.4	63.5	13.1	53.5	62.8	9.3	3.8
4	51.1	60.2	9.1	49.4	54.2	4.8	4.3
21	70.7	85.5	14.8	62.4	69.2	6.8	8
13	90	112	22	74.7	88	13.3	8.7
19	90	108	18	52.4	59	6.6	11.4
3	55	67.5	12.5	61.5	57.8	— 3.7	16.2
Av.	58.45	69.50	11.01	55.13	63.58	9.11	1.9
M			9.1			6.8	2.2
75%ile			13.52			12.67	5.22
25%ile			7.15			4.47	— 4.07
Q			3.18			4.1	4.65
P. E. (distribution)			3.81			4.19	6.1
P. E. (average)			$\pm 1.00$			$\pm 1.07$	$\pm 1.46$
						Av.=1.30 P. E.	
						M.=1.51 P. E.	

	6 months	12 months
Average of gains in test group in excess of control	1.37	1.9
Median	1.2	2.2
P. E. of difference	$\pm .48$	$\pm 1.46$
Average in terms of P. E.	2.85	1.30
Median in terms of P. E.	2.80	1.51

\*Numbers refer to cases as listed on Table II.

After a twelve months' interval, therefore, the actual average and median gains are slightly larger than after the first six months, but the variability is very much greater. Therefore, when expressed in terms of P. E., the gains are smaller. One of the test group cases (No. 13) who had gained 8 pounds after six months, gained 14 pounds in the second period of six months, making a total gain of 22 pounds. This gain is exceeded, however, by one in the control group (No. 12) who gained 3.5 pounds in six months, and 25.1 pounds more in the ensuing five months. This is certainly an enormous gain for five months, under any circumstances. Turning to Table XIV we find no corresponding gain in I. Q. for this child. Indeed there is a loss of five points.

Other children in the test group who made large gains, were case 12, with a gain of 18.5 pounds after twelve months, compared with 4.5 pounds after six months; case 19, gain of 6.3 pounds after first six months, and 18 pounds after 12 months; case 21, whose gain after the first period was 5.8 pounds, but who gained 14.8 pounds after twelve months. In these cases the gain in the second period greatly exceeds that for the first.

TABLE XVI  
Height, Second Retests, 13 Pairs

N*	Test Group (A)			Control Group (B)			A-B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
13	61.3	61.8	.5	56.8	60	3.2	-2.7
2	42.6	45.9	3.3	45.2	49.8	4.6	-1.3
28	54.9	56.3	1.4	53.4	55.9	2.5	-1.1
8	50.8	52.3	1.5	46.8	48.7	1.9	-.4
10	46.1	48.6	2.5	45.6	48.1	2.5	.0
1	46	49.7	3.7	46.4	49.7	3.3	.4
11	45.8	49.5	3.7	43.6	46.8	3.2	.5
16	51.6	53.5	1.9	48.7	50.1	1.4	.5
14	51	53.3	2.3	57.9	59.5	1.6	.7
3	47	50.2	3.2	51.7	54	2.3	.9
7	42.9	45.6	2.7	41.9	43.7	1.8	.9
4	47.5	50.5	3	48.9	50.8	1.9	1.1
19	57.7	60.5	2.8	46.7	4.8	1.3	1.5
Av.	49.63	52.17	2.5	48.74	51.16	2.42	.08
M			2.7			2.3	.5
75%ile			3.15			2.85	.85
25%ile			1.6			1.75	.92
Q			.775			.65	.885
P. E. (distribution)			.66			.78	.82
P. E. (average)			±.18			±.22	±.28
						Av.=	.29 P. E.
						M.=	1.79 P. E.

\*Numbers refer to cases as listed on Table II.



In weight, then, the mean gain of the test group over and above the control continues to increase through the second period of six months. The variability, however, increases enormously, which fact is due possibly to varying conditions which may enter in during the longer period to affect the health and thus lessen the gain of some of the children.

In order to determine whether the slight inequalities in interval length have any considerable effect on the results, we have calculated the relation between the length of interval and amount of improvement. The coefficient of correlation by the method of rank differences is equal to .03. The small number of cases renders the unreliability of correlation very great, but we can at least say that there is no consistent relationship between improvement and time interval, within the narrow limits here set. We are probably justified in taking twelve months as the interval, since such was the case in eight out of the fifteen test cases, while the greatest variation above this made was four months, and below it, one month.

The gains in height after twelve months are shown in Table XVI. The average gain of test group in excess of control, is only .08 inches, and the median .5 inches. Variability is about the same as at the end of six months, P. E.  $\pm$  .28. The average is only .29 P. E., but the median is a little larger, 1.79 P. E. If these measures are compared with the results after the first period, we have:

	6 months	12 months
Average of gains of test group in excess of control	.16	.08
Median of gains of test group in excess of control	.2	.5
P. E. of difference	.16	.28
Average in terms of P. E.	1.00	.29
Median in terms of P. E.	1.25	1.79

There seems to be little gain in height after the first period. Test cases 1 and 11 each show a gain of 3.7 inches after fifteen and twelve months respectively, compared with gains after six months of 1.6, and 1.9 inches. But case 2 in the control group, makes still greater comparative gain, +1.8 inches after six months and 4.6 inches after seventeen months. In this case there are almost six additional months for the child to grow, which may account for the larger gain. Control case 1, however, may be compared with his partner, mentioned above, since the interval between tests was the same for both. This boy grew 1.4 inches in six months, and 3.3 inches after 15 months. This is practically equal growth

with test case 1. Control case 11 also shows relatively great growth during 12 months, +3.2 inches, whereas the growth in six months was only 1.3 inches. Out of the test group, 7 cases gained more in the first period of six months, than in the second, while only 6 gained more in the second than in the first. Of the control group, 7 cases made more than half of their total gain during the second six months of the total twelve months' period. Since this is true, it seems likely that whatever increase in growth we find during the second half of the twelve months' interval, may be explained by incidental causes, and that so far as actual gain in height is considered, there is no further effect from the operations, after six months.

As was mentioned in the previous chapter, height and weight are of less significance when considered alone, than when taken in relation to each other and to the age of the individual. The gain in this weight-height-age relationship following upon operation for adenoids and tonsils, will be considered in the same manner as were weight and height gains. We have, then:

TABLE XVII  
Height-Weight Relationship, Second Retests, 13 Pairs  
Showing changes in per cent over or underweight after

N*	12 months' interval						A—B
	Test Group (A)		Change	Control Group (B)			
	Test 1	Test 3			Test 1	Test 2	Change
8	— 1	+ 7	+ 8	—13	0	+13	— 5
16	—16	—17	— 1	—13	— 9	+ 4	— 5
11	— 3	— 6	— 3	— 9	—10	— 1	— 2
14	—13	— 6	+ 7	— 5	+ 4	+ 9	— 2
19	+ 6	+10	+ 4	+ 1	+ 7	+ 6	— 2
28	— 7	— 4	+ 3	0	+ 5	+ 5	— 2
10	— 1	0	+ 1	— 6	— 6	0	+ 1
7	— 8	— 8	0	— 7	— 9	— 2	+ 2
2	— 8	— 4	+ 4	+ 9	+10	+ 1	+ 3
4	— 7	— 4	+ 3	—15	—15	0	+ 3
1	0	+ 6	+ 6	+ 8	+ 5	— 3	+ 9
13	—13	+ 6	+19	—11	—12	— 1	+20
3	+ 6	+13	+ 7	— 5	—20	—15	+22
Av.	— 5	— .54	4.46	— 5.08	3.85	1.23	3.23
M			4			—2	6
75%ile			6.75			2.75	3
25%ile			—2.50			2.75	— 2
Q			4.625			2.75	2.50
P. E. (distribution)			2.54			3.23	5.23
P. E. (average)			±.71			±.90	±1.15
						Av.=2.81	P. E.
						M.=5.22	P. E.

\*Numbers refer to cases as listed on Table II.

	6 months	12 months
Average of gains of test group in excess of control	1.83	3.23
Median of gains of test group in excess of control	4.00	6
P. E. of difference .....	.91	1.15
Average in terms of P. E. ....	2.02	2.81
Median in terms of P. E. ....	4.40	5.22

The mean of the actual gains in the second period exceeds that of the first. Again the second group of results is more variable, decreasing the reliability. There seems, however, to be a definite increase in the net gain of the test group during a second six months' period. Some individual cases may be cited. The greatest gain after six months is 8 units in the test case, matched by an equal gain of 8 units in the control group. After twelve months, the test group shows one gain of 19 units, the highest gain in the control group being 13. Six cases in the test group, and 13 in the control had lost at the end of six months, but after twelve months, all but 2 of the test cases showed a gain, and all but 5 of the controls. In 10 test cases out of the total 13, more than half of the gain occurred during the second six months. In the control group, six of the cases made more than half of their gain during the second six months, and the second interval gains of the other 7 cases exceeded the 50 per cent mark by so little that they may be accounted for by chance.

These results seem to indicate a slight but actual increase in the net gain of the test group during the second six months of the experiment, and an accompanying growth in the variability of these gains.

It will be remembered that the results described in the previous chapter show no gain in strength of grip as a result of operation. Comparison of the 13 cases tested after the second interval, with the 16 cases at the end of the first, gives results as follows:

TABLE XVIII  
Gain in Grip, Second Retest, 13 Pairs

N*	Test Group (A)			Control Group (B)			A-B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
7	7	6	-1	9	14	5	-6
19	22	20	-2	15	18	3	-5
2	9	10	1	10	15	5	-4
27	9	6	-3	8	9	1	-4
13	26.5	28.5	2	22	27	5	-3
10	12.5	12	-.5	10	12	2	-2.5
16	19	18.5	-.5	10	10	0	-.5



TABLE XVIII (Continued)

N*	Test Group (A)			Control Group (B)			A-B
	Test 1	Test 2	Gain	Test 1	Test 2	Gain	
3	12.5	13	.5	14	14.5	.5	0
14	16	17	1	22	22.5	.5	.5
21	16.5	15	-1.5	19	17	-2	.5
1	13	18	5	11	15	4	1
23	11.5	15	3.5	14.5	16	1.5	2
11	11	11	0	11.5	7.5	-4	4
Av.	14.27	14.61	.34	13.54	15.19	1.65	-1.31
M			0			1.5	-1.5
75%ile			1.25			3.75	.62
25%ile			-1.38			.12	-.4
Q			1.31			1.81	.51
P. E. (distribution)			1.34			1.65	2.31
P. E. (average)			±.37			±.46	±.59
						Av.=	-2.22 P. E.
						M.=	-2.54 P. E.

	6 months	12 months
Average of gains of test group in excess of control	— .24	—1.31
Median of gains of test group in excess of control	—1.00	—1.50
P. E. of difference	± .48	± .59
Average in terms of P. E.	— .50	—2.22
Median in terms of P. E.	—2.08	—2.54

The greatest gain in the test group after twelve months is 5 Kg. (Case 1). During the first six months this case lost 1 Kg. There are two gains of 5 Kg. in the control group. Of these two (cases 7 and 13) had gained 1 Kg. during the first interval and another (case 2) 2 Kg. The greatest loss in the test group after the twelve months' period was 3 Kg., by case 27, which had already lost this amount at the end of six months. The greatest loss in the control group was suffered by case 11, a loss of 4 Kg., all in the second period. After the first period, 9 out of 16 cases in the test group gained in strength of grip, and 13 in the control group. After the second period, the test cases showing gain numbered only 7 out of 13, while all of the control cases had gained except 2. Of the test group 8 cases in the second period either gained less than half of the amount they had improved in the first period, or dropped from the scores they had made at that time. The corresponding numbers for the control group are 6 and 7.

There is evidently no improvement in strength of grip twelve months after operation. The unreliability of the results is very great. However, there is certainly no tendency toward improvement. Why this should be is a question. It

\*Numbers refer to cases as listed on Table II.

may be that the change in examiners is partly responsible, for performance in this test is influenced to a surprising extent by the manner in which it is presented.

TABLE XIX  
Tapping—Second Retest, 15 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
21	152	154	2	149	174	34	—32
13	160	176	16	150	188	38	—22
1	135	142	7	106	134	28	—21
18	133	126	— 7	100	108	8	—15
23	150	157	7	122	141	19	—12
10	68	145	77	70	148	78	— 1
3	136	144	8	135	135	0	8
20	105	122	17	150	154	4	13
15	172	192	20	140	145	5	15
27	108	114	6	115	101	—14	20
8	113	128	15	131	121	—10	25
28	150	176	26	178	172	— 6	32
2	105	135	30	152	139	—13	43
11	125	120	— 5	155	102	—53	48
14	190	228	38	175	165	—10	48
Av.	133.47	150.6	17.13	134.6	141.8	7.2	9.93
M			15			4	11
75%ile			21.5			21.5	26.75
25%ile			5			—10.75	—16.5
Q			8.25			16.12	21.62
P. E. (distribution)			10.13			17.2	22.07
P. E. (average)			±2.67			±4.53	±5.26
						Av.=1.89	P. E.
						M.=2.09	P. E.

There were 15 pairs of cases who performed the tapping test at the end of twelve months. Comparison with the 21 pairs after six months yields the following results:

	6 months	12 months
Average of gains of test group in excess of control	— .09	9.93
Median of gains of test group in excess of control	0	11
P. E. of difference	±3.10	±5.26
Average in terms of P. E.	.03	1.89
Median in terms of P. E.	0	2.09

The gain in the second interval is greater for the tapping test than for any of the tests yet described. After the first six months there is no gain. At the end of ten months the average gain is 9.93, and the median 11 taps per half minute. After six months' interval, 11 of 21 test group cases had lost. At the end of twelve months, only 2 out of 15 had lost. The

\*Numbers refer to cases as listed on Table II.

control group, on the other hand, lost in 11 out of 21 cases after six months, and in 6 out of 15 at the end of twelve months. All but one of the test group cases made more than half of their gain in the second period. Of the control group only 7 cases did this. The variability of gains after 12 months is about equal to the variability at the end of six months.

Strangely enough, decrease in fatigueability as described in the previous chapter does not show itself after 12 months. In fact, the negligible loss in ability noticeable after six months has increased after a period of twelve months. In only 4 out of 11 test group cases, is the gain in the second period equal to that of the first, a similar result to that found in the control group, where 5 out of the 11 cases made half their total gain in the second interval. The results are compared below.

	6 months	12 months
Average of gains of test group in excess of control	— .020	— .06
Median of gains of test group in excess of control	— .015	— .09
P. E. of difference	± .04	± .036
Average in terms of P. E.	— .50	—1.66
Median in terms of P. E.	— .38	—2.50

TABLE XX  
Tapping for Fatigue, Second Retests, 11 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
20	— .10	.03	— .13	.20	.0	.20	— .33
27	.15	.17	— .02	.09	— .17	.26	— .28
10	— .09	.27	— .36	.06	.16	— .10	— .26
14	.09	.05	.04	.13	— .13	.26	— .22
13	— .03	— .06	.03	.27	.07	.20	— .17
28	.01	.05	— .04	.17	.10	.07	— .11
15	.03	.03	.0	.18	.08	.10	— .10
23	.11	.0	.11	.06	.0	.06	.05
21	.27	— .01	.28	.29	.14	.15	.13
11	.28	— .04	.32	.11	— .01	.12	.20
18	.14	— .15	.29	.01	.15	— .14	.43
Av.	.078	.031	.047	.143	.035	.107	— .06
M			.03			.12	— .09
75%ile			.153			.20	.07
25%ile			— .062			.02	— .275
Q			.107			.09	.172
P. E. (distribution)			.087			.093	.19
P. E. (average)			±.02			±.03	±.036
						Av. = —1.66 P. E.	
						M. = —2.50 P. E.	

\*Numbers refer to cases as listed on Table II.



The point of greatest interest in the present study is, as has been said, improvement in intelligence. Does operation for adenoids and tonsils result in improvement in intelligence, as measured by I. Q.? If such improvement does not manifest itself after six months, can it be found after a second period of the same length? The latter question is answered by observation of Table XXI and attention to the following facts, gathered from the 21 pairs of cases who were given intelligence tests after the twelve months' interval.

	6 months	12 months
Average of gains of test group in excess of control	—1.035	—3.14
Median of gains of test group in excess of control	—1	—3
P. E. of difference .....	±1.10	±1.84
Average in terms of P. E. ....	— .94	—1.71
Median in terms of P. E. ....	— .99	—1.63

The result after twelve months remains the same as that after the six months' interval. A gain or loss of two or three points in I. Q. is negligible, so that the mean gain of the test group in excess of the control is practically zero at the end of each period. Variability increases with the length of the interval. One case in the test group (case 4) gained nothing in six months, but showed a gain of 15 points after 13 months. However, there is a control case to match this,—case 15, who gained 1 point in the first six months and 20 points after 10 months. Case 20 in the test group lost 4 points in the first six months, but gained back these and 10 additional in the second period. But control No. 1 gained 5 points in the second interval after having lost 4 in the first.

On the other hand several cases lost in the second period, as compared with the first. Test group case 7, for example, gained 8 points in the first six months, and lost 6 of them in the second. Case 3 in the same group lost 3 points in the first period, and failed to regain any of them. Case 12 lost 3 points in six months and 9 more before the end of 12 months. In the control group, case 23 gained 15 pounds in the first six months and lost eight of them in the second. Summing up gains and losses in the second period, for both groups:

TABLE XXI

I. Q., Second Rereests, 21 Pairs

N*	Test Group (A)			Control Group (B)			A-B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
27	110	112	2	72	98	26	-24
15	97	97	0	78	98	20	-20
3	94	91	-3	82	96	14	-17
7	91	93	2	95	112	17	-15
20	106	116	10	118	140	22	-12
23	85	80	-5	81	88	7	-12
8	91	92	1	98	111	13	-12
12	98	86	-12	98	91	-7	-5
21	64	66	2	86	93	7	-5
4	96	111	15	83	102	19	-4
13	70	61	-9	66	60	-6	-3
18	98	101	3	90	92	2	1
16	65	74	9	74	81	7	2
10	110	116	6	104	107	3	3
28	81	83	2	95	94	-1	3
11	103	102	-1	101	95	-6	5
1	82	93	11	80	85	5	6
14	96	102	6	140	138	-2	8
19	96	97	1	98	90	-8	9
2	107	113	6	91	86	-5	11
9	83	102	19	87	91	4	15
Av.	91.5	94.6	3.0	91.2	97.5	6.2	-3.1
M			2			5	-3
75%ile			6			13.75	4.5
25%ile			-7.5			-4.25	-12
Q			3.37			9.00	8.25
P. E. (distribution)			2.09			8.24	8.86
P. E. (average)			±.45			±1.79	±1.84
							Av.—1.71 P. E.
							M.—1.63 P. E.

	Lost in 2nd period	Gained in 2nd period	No change	Gained equally with gain in 1st period
Test group	9	11	1	9
Control group	10	11		8

The average gain of the test group was 3.09 points after 12 months, compared with 2.25 after six. The control group, however, made an average gain of 6.24 after 12 months, the gain after six months being 3.29. These numbers are insignificant as gains, but they at least show no improvement in the test group which the control group does not reveal as well. On the basis of the results, we may say that there has been no improvement in I. Q. as a result of operation, either after six months or after twelve.

There remains to be considered only the result of the Healy Picture Completion Test. We have scores in this test for 18 pairs of cases. When these scores are compared with those in the former tests, the results stand as follows:

\*Numbers refer to cases as listed on Table II.

Average of gains of test group in excess of control	—5.85	—3.36
Median of gains of test group in excess of control	—7.75	1
P. E. of difference	±3.05	±3.38
Average in terms of P. E.	—1.92	— .87
Median in terms of P. E.	—2.54	.26

The figures given above show no gain in the Healy test as a result of operation. Both after six months, and after twelve, we find that the test group has gained no more than the control group. As before there are individual cases showing considerable gain in the second period, but these are matched by control cases which reveal equal or even greater gains. In the test group, 5 cases lost in the second period in comparison with the first, 18 gained, and 7 gained as much in the second period as in the first. The control group lost in the second period in 5 cases, gained in 18, and gained as much as in the first period in 7 cases. The two groups, then, are practically equal, both showing a gain in the second period, but this gain cannot be due to the operations, since the control group did not undergo operation.

TABLE XXII

Healy A, Second Retests, 18 Pairs

N*	Test Group (A)			Control Group (B)			A—B
	Test 1	Test 3	Gain	Test 1	Test 3	Gain	
12	20	1	—19	—10	41.5	51.5	—70.5
23	49.5	62.5	13	4	64	60	—47
15	7	19	12	1	54.5	53.5	—41.5
1	—22	11	36	—50	11	61	—25
2	28.5	24.5	—4	3	19.5	16.5	—20.5
3	21.5	15.5	—6	17	25.5	8.5	—14.5
19	57.5	60.5	3	—22	—15	7	—4
13	43	62.5	19.5	—1.5	21.5	23	—3.5
21	20	38	18	70.5	88	17.5	.5
10	—12	11	23	27	48.5	21.5	1.5
14	12.5	77	64.5	—5	48.5	53.5	11
27	—25	25.5	50.5	2	39.5	37.5	13
11	—8	15.5	23.5	—29	—20	9	14.5
16	49	79	30	30	45.5	15.5	14.5
7	—6	16	22	45	11	6.5	15.5
28	29.5	77.5	48	64.5	83.5	19	29
18	—13.5	13.5	27	—32	—35	—3	30
20	0	55	55	30	48.5	18.5	36.5
Av.	13.81	36.92	23.11	5.77	32.25	26.47	—3.36
M			22.5			18.75	1
75%ile			33			44	14.5
25%ile			7.5			8.75	—22.25
Q			12.75			17.62	18.37
P. E. (distribution)			11.5			17.72	17.86
P. E. (average)			±2.7			±4.22	±3.88
						Av. = —	.87 P. E.
						M. = +	.26 P. E.

\*Numbers refer to cases as listed on Table II.



TABLE XXIII

Gains	Weight in pounds		Height in inches		Height-weight percents		Grip in Kg.	
	lb.	P. E.	in.	P. E.	%	P. E.	Kg.	P. E.
Av. 1	1.37	2.85	.16	1.	1.83	2.02	— .24	— .50
Av. 2	1.90	1.30	.08	.29	3.23	2.81	— 1.31	— 2.22
M. 1	1.20	2.80	.20	1.25	4.00	4.40	— 1.00	— 2.03
M. 2	2.20	1.51	.50	1.79	6.00	5.22	— 1.50	— 2.54
P. E. 1	± .48		± .16		± .91		± .48	
P. E. 2	± 1.46		± .28		± 1.15		± .59	

TABLE XXIII (Continued)

Gains	Taps in ½ min.		Tapping in fatigue ratios		I. Q.		Healy Score	
	taps	P. E.	ratio	P. E.	Pts.	P. E.	Pts.	P. E.
Av. 1	.09	.03	— .02	— .50	— 1.035	— .94	— 5.85	— .92
Av. 2	9.93	1.89	— .065	— 1.66	— 3.14	— 1.71	— 3.36	— .87
M. 1	0.	0.	— .015	— .38	— 1.	— .99	— 7.75	— 2.54
M. 2	11.	2.09	— .09	— 2.50	— 3.	— 1.63	1	.26
P. E. 1	± 3.10		± .04		± 1.10		± 3.05	
P. E. 2	± 5.26		± .036		± 1.84		± 3.88	

In Table XXIII are collected the results discussed in the foregoing chapter. The mean results are expressed in terms of P. E. and as gross values, so that the various tests may be compared.

## CHAPTER V.

### SUMMARY.

The results obtained from the experiment may be summarized as follows:

1. Six months after operation for adenoids and tonsils, there seems to be a slight but not very reliable gain in weight as the result of the operation. After twelve months this has increased; indeed, it has very nearly doubled.

2. Gain in height, resulting from operation, is so slight as to be unreliable. This gain does not increase during a second period of six months.

3. The height-weight-age relationship is an excellent measure of the physical well-being of the child. The figures expressing this relationship show no very reliable gain in the first six months, but improvement increases considerably during the second period.

4. The test group shows no gain over the control group in strength of grip. There seems on the other hand to be a slight loss; which does not decrease in amount during the second period.

5. Speed of tapping did not increase during the first period, any more for the test group than for the control. During the second period, however, there is a marked improvement.

6. Operation for adenoids and tonsils does not lessen fatigueability as shown by the tapping test. The probability is, however, that the test is at fault.

7. No rise in I. Q., as a result of operation, makes itself evident after six months or after twelve months.

8. There is no improvement in the performance of the Healy test either after six months or after twelve.

9. In every test except grip and tapping, there is a marked increase in the variability of the gains after the second period. This is possibly due simply to the fact that the longer interval

permits the intervention of more extraneous factors which may influence the scores in one direction or the other.

10. A group of 236 children with diseased tonsils showed equal distribution of I. Q. with a group of 294 children who were normal in this respect.

## CONCLUSIONS

The article by MacPhail, which has been reviewed in a former chapter, showed pretty conclusively that the removal of adenoids and tonsils was followed by improvement in school work. That such improvement was not due to a rise in general intelligence can be concluded from the present experiment. That efficiency in school work does not rest wholly upon intelligence has been demonstrated more than once. The tendency here noted to improve in general physical tone may, perhaps, serve as a sign of the factor upon which such improvement depends. Improved health means better attention, better emotional response, greater resistance to fatigue, and probably increased efficiency.

Interesting investigations of such improvement in efficiency might be made by administering educational tests to groups similar to those of the present study. The results of such an experiment would be exceedingly instructive, and would be more significant than conclusions drawn from school marks. By this means also we might determine along what special line efficiency is most affected.

Since there was no recuperation in intelligence resulting from operation for adenoids and tonsils, it is reasonable to expect that there had been no retardation from which to recuperate. This supposition is borne out by results of the statistical study, wherein we found that a group of children suffering from diseased tonsils possessed equal intelligence with a group which was free from such defect.

We can say to physicians, then, with fair amount of assurance, that removal of adenoids and tonsils will probably not raise to any great degree the intelligence level of the mentally defective child who is brought to him. We can say to students of the constancy of the I. Q., that it is not greatly lowered by adenoids and diseased tonsils and we may say to the clinical psychologist that these defects have no demonstrable effect upon general intelligence, whatever effects they may have on



volitional and emotional normality,—the two elements which, along with intelligence are necessary for the maintenance of the individual as an instrument of social efficiency.

# AN EXPERIMENTAL STUDY OF THE FACTORS AND TYPES OF VOLUNTARY CHOICE

BY  
ALFRED H. MARTIN

ARCHIVES OF PSYCHOLOGY  
EDITED BY R. S. WOODWORTH  
No. 51

NEW YORK  
MAY, 1922

226736<sup>d</sup>  
7. 11. 28





## CONTENTS

I. The Problem and its History .....	5
(a) Theoretical Discussion.	
(b) Experimental Results.	
II. Methods and Procedure .....	12
First Series.	
Second Series.	
III. Qualitative Analysis .....	23
General Description.	
(a) Fore Period.	
(b) Main Period.	
(c) Reversal Period.	
(d) Motivation.	
IV. Statistical Results .....	89
Chances of Reversal.	
Time.	
Difficulty.	
Confidence.	
V. Discussion of Results .....	104
VI. Summary .....	113
VII. Bibliography .....	115



# An Experimental Study of the Factors and Types of Voluntary Choice

---

## I.

### INTRODUCTION AND HISTORY.

The problem of will is twofold in that it is inclusive of "inner and outer volition." While the latter is single, the former implies the latter as a consummation of the internal act. "Inner volition" implies deliberation and choice, resulting in a decision to carry out the chosen course of action, "outer" involves a resolution to carry out the act either selected or obviously present to the mind. The former process is, of its very nature, intellectual, with a final resolution into a "mental set," while in the latter case the "set" is arrived at without a deliberate act of choice. The main purpose of this study is an attempt, first at verification of the work already carried out in the field of voluntary choice, and second at a further extension of the investigation to the nature of the mental set involved.

More than in any other field of the thought processes, speculative and theoretical considerations have been put forward, that have shaped and guided investigation, so that even in a very brief survey of the history of the subject a treatment of this aspect, no less than of the strictly experimental side, would appear to be necessary. Further, experimental investigation is meagre and limited, and is supplemented in many textual expositions by such general retrospective speculation.

#### (a). Theoretical Discussion.

James (9. Chap. XXVI.) regards the basis of will as constituted by an inherited tendency to action, and, with the arousal of more than one such tendency, a plurality of possibilities in the shape of impulses, is generated; the result is a mutual inhibition or suspension of action until one of them prevails. When the thought processes enter in as an integral part, the phenomenon of voluntary choice may be said to be



present. Five distinct types of decision are recognized by him and consist of,

- (a). The "reasonable type," where the arguments naturally settle themselves and the process is largely rational.
- (b). Where the decision is "given in haste and without due deliberation," owing to the immediate pressure of circumstances.
- (c). The type where "the individual wearies of the conflict" and commits himself to one or other alternative.
- (d). "From the easy and careless to the sober and strenuous level," due to the introduction of some moral factor.
- (e). The type where decision is accompanied by a "sense of effort," where we feel that we have deliberately chosen.

Stout (19. Chap. X.) uses four similar categories with almost equivalent terminology. Calkins (5. Chap. XII.) in considering these types, divides them into two main classes, first, those where a sense of effort is present, that is to say, where we appear to act against the line of greatest resistance and second, those where an established system of comparative values makes the choice a mere imaginal comparison of such values. In the former case, in spite of the effort of choice, the two alternatives still persist though decision in favour of one of them has been arrived at, while in the latter case the resulting act excludes the one possibility with the acceptance of the other alternative. Wundt (26. No. 14.) likewise differentiates the two types, by the use of different terms for the resulting mental sets. Where marked conflict is present with a state of doubt which is only ended by a strong effort, then the process is one of "decision," which is intensively stronger in consequence than when the process is comparatively easy and ends in "resolution."

The consideration arises, as to what is the main factor in the "sense of effort" and in other types of decisions. James suggests that it comes from a greater degree of attention to one factor rather than the other, but does not explain whence the additional degree of attention originates. McDougall (11. pp. 242 and 247.) follows James in regard to his identification of the "effort" with an increased attention as the factor, but adds, as an explanation, that, "in this self-conscious act of will, the idea of self or self-consciousness is able to play its great role in volition only in virtue of the self-regarding

sentiment, the system of emotional and conative dispositions that is organized about the idea of the self and is always brought into play to some extent when the idea of the self rises to the focus of consciousness. The conations, the desires and aversions arising within this self-regarding sentiment are the motive forces which, adding themselves to the weaker ideal motive in the case of moral effort, enable it to win the mastery over some stronger, coarser desire of our primitive animal nature and to banish from consciousness the idea of the end of this desire."

Stout (19. p. 707.) treats the problem in similar fashion, but where McDougall definitely attributes "sense of effort" only to the presence of the self-regarding sentiment, Stout refers all forms of volitional choice to the interposition of a self-factor, and "under the concept self as expressed in the word "I" is included in systematic unity the life history of the individual, past, present and future, as it appears to himself and others together with all its possible or imaginary developments" . . . . "The motives are motives only in so far as they arise from the nature of the self, and presuppose the connection with the "self" as a determining factor. From this it follows that the recognized reasons for a decision can never constitute the entire cause of a decision. Behind them there always lies the idea of the self as a whole. What this involves can never be completely analysed or stated in the form of definite reasons or special motives." Calkins (5 p. 226) also specifically emphasizes the ego-centric nature of volition. "In will I am actively, assertively, related to my environment, I am conscious of my superiority and my independence of it, I conceive of it as existing mainly for my own use or gratification." And she finally dissents emphatically from any attempt to define the act merely in structural terms of sensation, image and feeling, the "consciousness of the self as willing" being utterly irreducible to such.

Pillsbury (16. p. 525.) summarises the action of voluntary decision as "the whole man active in coming to the point." Warren (22. pp. 310-311. defines volition as selective "in that it tends to bring about the fittest actions." It is marked by anticipatory images from which this selection takes place. When a real decision is reached, the image becomes a purpose idea. The purpose image then is based upon a memory image aroused by distant sense data; and the distinguishing mark

between a purpose idea and other anticipation images is the prominence of our own activity in the purpose experience. Purposes are thus anticipatory representations of our actions, and they approach more or less to the rational type of behaviour.

Ogden (15. p. 172.), while postulating a growing purposiveness and polarisation of self and object, finds in volition the "consciousness of a self-determined act". . . . "such as does not appear in any other form of mental happening." Like Calkins, while acknowledging the presence of the structural elements, he regards the consciousness of the "actual moment of self" as a distinct advance in psychological discovery. Titchener (20. p. 467) also recognizes that in the will act, in which voluntary decision will be included as a type, there is essentially a reference to the psychological self, but would insist "that the will attitude is rarely explicit or focal but is, as a rule, incorporated in the total reaction consciousness." Külpe's (10. p. 265.) position is opposed to any such factor of decision as the "self" entity; his attitude is purely "structural" in character. The "self" that operates in the formulation of a decision is only an interpretative shadow whose true origin is in the bodily self, and when a total response of the latter by means of a kinaesthetic image or sensation is made at the temporal point of decision, then this effect is transformed into a more or less philosophical but unpsychological representation of self. As to the selective principle which dominates the procedure and maintains direction to a final goal, this is, for him, amply provided for by Wundt's principle of apperception.

To summarize the preceding, voluntary acts differ from impulsive types in the fact that deliberation precedes the consummation of the act. This deliberation proceeds upon a basis of the individual's past experience and personal evaluation of the alternatives. In the final act of choice the self identifies itself with the act, and the voluntary choice or decision becomes complete in itself. In certain cases of conflict the decision is only ended by a distinct "effort" of the subject, hence the "ego" factor is increasingly evident as a moment of decision; this type is characteristically the unique type of decision. Other types, such as forced decisions, may be arrived at from pressure of time, or circumstances, such as lack of knowledge; or one factor of greater weight may so exceed the others in importance as to be accepted without conflict. The latter are



variable types that approach, but do not attain to, the highest levels of voluntary choice.

(b). Experimental Results.

The history of the experimentation on the side of volition begins with Ach (1), who tested out Wundt's logical analysis by experimentation and discovered that the "will-act" really took place in the fore period through the acceptance of the Aufgabe of the directions as understood by the subject. During the main period, the will act was consummated according to the previous "mental set" achieved by the subject.

His next researches (2) led him into further investigation of the will act by the interference method. He caused his subjects to learn nonsense syllables of three letters by the method of "paired associates," the strength of such associations being measured, in psychological terms, by the number of repetitions. After the learning part was completed, the subject was given the task of responding to a list which included the original stimulus words under the special direction of either riming, reversing the first and last letters, or using an unrelated word. As, in some cases, the preformed associations interfered with the performance of the will act, this resulting disturbance was measurable by means of the chronoscope in temporal terms. When the subject became conscious of his error, the result was a strengthening of the "determining tendency." Introspection showed that this was due to the inner resolve to do better when the internal distraction was recognized. This "determining tendency" Ach ascribes as directly due to the interposition of the self entering in as a conscious factor, "I can and I will" being its internally vocal or ideational expression.

The same year, Rowe, (17), investigating the voluntary process on the side of development of motor skill by means of typewriting methods and unfamiliar writing postures, found that the methods adopted were at first perceptual and ideational, but that these "gropings" became automatised in a definite kinaesthetic set which was of an ideo-motor character.

These investigations of the will act itself, however, do not afford any light upon the matter of voluntary choice, and were criticized on this account by Selz (18), in his appraisal of Ach's work. The investigation of this side of the will act was taken up by Michotte and Prüm (13). Seeking the simplest

material possible, they used the operations of addition or subtraction, or of multiplication or division of simple numbers, and caused their subjects to choose the working out of one of the two operations for "weighty reasons." One of the pairs of the alternatives was shown in the fore period, while the numbers were exhibited in the main period when the choice took place, the stimuli being exposed by means of a card-changing apparatus. Their results show that the process divides into well-defined and articulated parts.

- (1) During the fore period comes:
  - (a) The acceptance of the Aufgabe,
  - (b) A subsequent period of expectancy.
- (2) In the main period:
  - (a) The realization of alternatives in the case,
  - (b) The consideration of these alternatives,
  - (c) The act of choice.

The investigation was completed before the publication of Ach's work, and independent of it, but is corroborative of it in regard to the actual "moment of choice." This factor designated as "*la conscience d'effort*" the authors identify with that of Ach's determining tendency, i.e. that of the self. Michotte, in a supplementary note to the original work, makes a comparison of his own work with that of Ach (2) which had appeared just previous to his own publication. Criticism may be directed to the work from two points, first, that the choice is not finally consummated into a will act, and again, that the motives are too trifling for such weighty process as voluntary choice. Such criticism the authors have anticipated, and reply that the pressing of the reaction key really appeared as the consummation of the act, and in the second regard the subjects felt as if they were really making an important decision.

Barret (3) carried out investigations in motivation, using as a stimulus eight beverages, ranging from unpleasant to pleasant in taste and distinguished by names. When the association of name and taste was complete, they were used in regular series. The time of choice and the time of lifting the glass containing the drink was noted, and there was found a regular decrease in reaction times of both; the values, already quite familiar singly, being soon learned relatively between each pair of the series. His work was not confirmatory to that of Ach or Michotte and Prüm, principally, as Michotte

points out, on account of the automatised form of the process of choice.

In a study of decision types by Bridges, (5) investigating the relation between the constancy, time and accuracy of decision, a variety of material was employed, such as the arrangement of series of names of persons and objects. The same experiments were repeated after given periods of time had elapsed to test the subject's consistency. The term "decision" however, does not appear to be a happy one, for the process usually so designated is specifically related to volition, while, so far as the work of Bridges shows, the choice was probably made for aesthetic reasons rather than for the purpose of future activity, and should pertain to the category of judgment.

Wheeler (24) followed up the attempt to improve the methods (a) by using more important and concrete practical material, such as pictures and pianoforte selections, and, (b) by actually fulfilling the act of choice. The names of two pianoforte selections were exposed for the subject to choose from, the piece chosen being played afterwards on a pianola. Again, two pictures were exposed, the subject choosing one of them to hang in his room. Results confirmatory of those of Michotte and Prüm's work were obtained, so far as the structural elements are concerned, but the "decision factor" of the self was pronounced to be lacking, the "self moment" that appeared in previous work being due, the writer thinks, to incomplete analysis, lack of training in the reagent, or individual prejudice in the interpretation of the results on the part of the experimenter. Ultimate analysis must yield only elementary sensory and imaginal experiences, such a term as "consciousness of self" being synthetic and interpretative rather than final and analytic. Wheeler's position is typically structural in character, and his analysis closely follows that given by Külpe.

The result of the experimental work, then, is the discovery of the fact that, when a mental set originated by an act of will is interfered with, the result of such disturbance is to arouse an added measure of the original determining tendency to overcome it. In the case of voluntary choice, the act of decision is also effected by this identical factor, viz. the self that functions as a determining tendency. The validity of such an "interpretative functioning" has been challenged by the contention that all such entities may be reduced to analytic structural factors of sensation, image and feeling.



## II. METHODS AND PROCEDURE

The aim of the present investigation was twofold; first to verify the previous work done upon voluntary choice, and second to examine the after period by interference with choice by any available methods. The work was carried out in the graduate Laboratory of Columbia University, during the two semesters of the year 1920-1921, under the direction of Professor Woodworth.

In the first series the subject was asked, during the fore period, to imagine himself in a certain predicament, and then, during the main period to consider two alternatives in order to solve his dilemma and make a choice between them. This having been done, full introspections were then taken. Where necessary, questioning was resorted to in order to clear up obscurities, after the method of Ach, but this became less and less necessary with the progress of the subject in the work. A certain evaluative judgment of the subject's confidence in his choice, and its degree of seriousness and difficulty, were also obtained. In certain of the situations the subject was next required to "reconsider his decision, with a view to its reversal." Further introspections of this stage were then taken, and the degree of possibility of reversal also noted. Since these situations were more or less hypothetical and the decisions impossible of immediate execution, the work was checked up by a second series wherein the reagent was required to choose one of two odors, the one chosen being smelled by him for a period of two seconds, thus actually carrying the decision into execution. The procedure was similar to that of the first series. The time in every case was taken by means of stop-watches marking fifths of seconds. As the time varied from 2.5" to 190", such a means presented a sufficiently accurate method of time record for the purpose of the experiment.

The subjects consisted of two members of the Teaching Staff of Columbia, in the Department of Psychology, designated S. I. and S. II., four graduate students in Psychology, of whom S. III. and S. IV. were men and S. V. and S. VI. were women, and two undergraduate women students, S. VII. and

S. VIII., who had attended courses and engaged in laboratory work in psychology. To all these the experimenter desires to express his gratitude for their cheerful and ready endurance, both of protracted sittings and of frequently disturbing experiences.

Since the process was decidedly mental the image types of these subjects are incidentally of interest. They were predominantly visual with kinaesthetic forms present to a less degree; the single exception was one of the graduate women,—S. V. whose verbal-motor speech images were even more frequent than the visual. This subject, with one member of the Teaching Staff, S. I., who also had marked kinaesthetic images, exhibited a peculiar tendency to schematise, the one in visual, the other in visual and verbal motor terms; this was also noted in others, but in a much less marked degree. Thus S. I.'s reference to "short, acute pains" as compared with "slight protracted pains" visually appeared as two contrasted normal curves of distribution. (Fig. 1.)

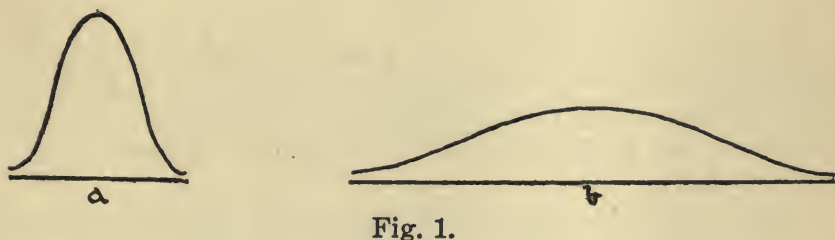


Fig. 1.

In a second case, one of synaesthesia, of the same subject in the fore period (a) of No. 28. . . . "There developed a vague visual schema, (Fig. 2.) with certain kinaesthetic elements also present. On the left was a line and a movement rather long, and apprehended as vertical. On the right was a shorter line of that kind, and present in both was a kinaesthetic experience of a feeling of length, as if I looked along each and realised the length of one and the shortness of the other. With the short line on the right was a kinaesthetic reference from that line outward, and I was aware that this represented the relation of this girl, whom the short line stood for, to her father. The father himself was not represented. Apart from the schematic representations of the relation the process was imageless."

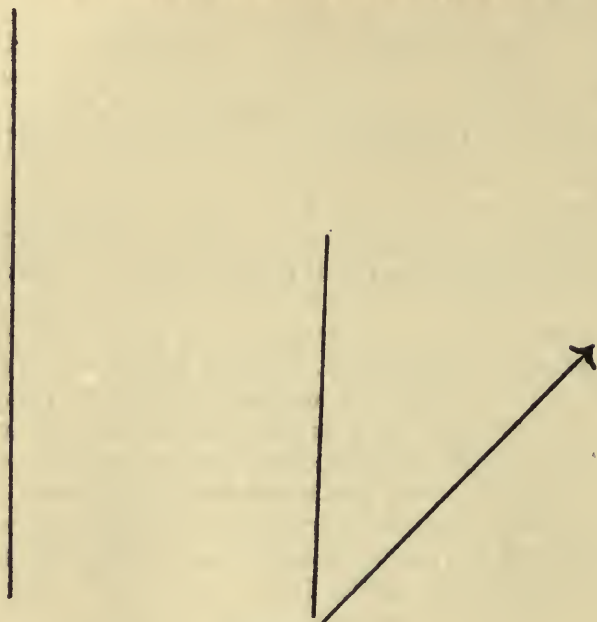


Fig. 2.

In regard to the latter, S. V., whose schematic references, also synaesthetic, were less frequent, the following is typical. The reference comes from the forepart (a) of problem No. 21. "Reading the situation brought the following scheme (Fig. 3.) into my mind."

"Eye movements were experienced as the eyes seemed to travel along the lines, while in the background were two vague pictures of Edison, one as a lad of twelve and one at the present time. The lines to the left were hazy and indistinct and seemed to lead off into space, since I could not remember other mechanical achievements, but the others led up to more or less clear visual images of the actual objects."

The first series consisted of problems that contained (a) a situation and (b) the alternatives that offered a solution of the situation. There were thirty of these in all, but included among them were a certain number of judgments that could by no means be classed as decisions but were prepared in similar fashion to the others. The situations and alternatives were suggested by a class of graduate students as being typical of those in which they had actually found themselves. Out



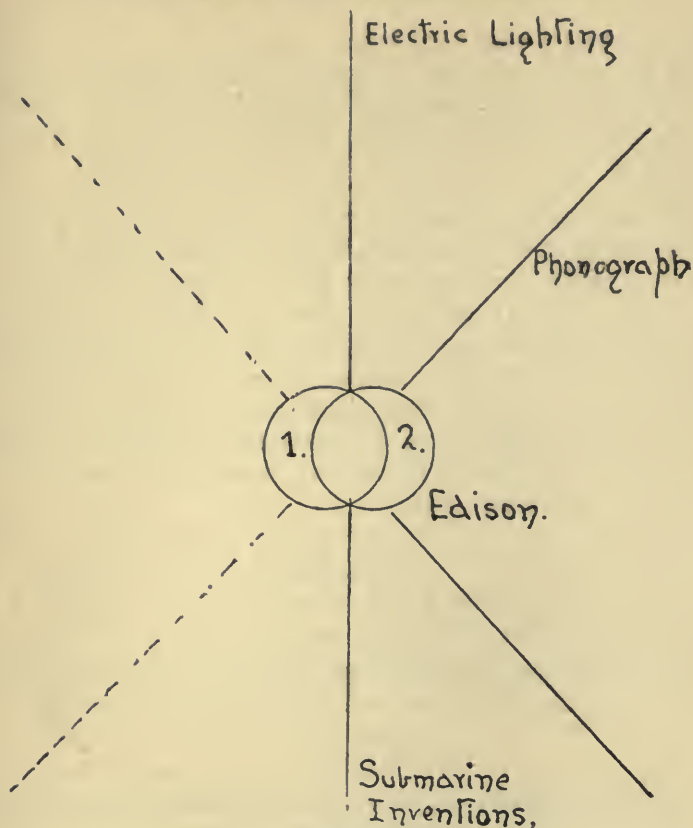


Fig. 3.

of the mass handed in, the members of the series used were selected as being reasonable, practical and fairly general in type. They were then each arranged in two parts, constituting the fore and the main periods, as previously mentioned, each part being separately typewritten on a slip of paper. In addition, pictures were used in five cases, these appearing as alternatives in the main period. They were arranged in grey surrounds and pasted side by side on stiff card. These picture choices, situations and inferences or judgments were combined, and comprised in all thirty problems in this series. They were graded in regard to the combined degree of difficulty and seriousness of the problem into four classes, and an average obtained by means of this classification. The judges, ten in number, were members of the Staff of the Departments of Psychology within the University. While the variability

of the judgments was very great, plainly indicating the presence of subjective factors, at least a series that approximated a progressive arrangement in these two respects was obtained. The "judgments" were included with the "decisions" and given positions with the others as if of a similar type.

Below are typical examples; where "R" is marked after the number which indicates place in the series, the subject was asked to consider the possibility of a reversal of his decision. The numbers where this was required were indicated to the experimenter, however, by a private mark on the covering page, namely the inclusion or omission of the punctuation point after the number, so that the subject was not aware of the time when this requirement was forthcoming. Half of the examples, fifteen in number, carried this requirement but the succession was in chance order.

No. 3. R.

- (a). You are attending a match and your side appears to be on the point of winning,
- (b). Someone offers to bet a box of candy that the opposing side will be victorious; would you accept or reject?

No. 6. R.

- (a). The hour is late and the day has been tiring; you are about to take the "subway" home, when suddenly you discover you have lost your money;
- (b). Would you choose to walk the distance home or risk explaining your quandary to some decent looking stranger?

No. 11. R.

- (a). You are preparing to attend a social gathering at a home not previously visited.
- (b). Would you prefer to go in evening dress with a chance of being made conspicuous, or in ordinary dress and perhaps feel out of place?

No. 12.

- (a). You are convalescing; it is your first day out of doors and the weather is bright but cool.
- (b). Would you prefer to sit in the sun in an uncomfortable chair or in the shade in a comfortable one?

## No. 13.

- (a). You are telling a story to an interested company, when suddenly you realize it has a personal bearing on one of them.
- (b). Would you choose to disappoint your hearers by suddenly breaking off, or continue in deference to your audience regardless of the individual's feelings?

## No. 15.

- (a). You will see two coloured reproductions of Landscapes by Van Hier.
- (b). Choose which of these you would prefer to hang in your own room.

## No. 19. R.

- (a). On returning home after making a purchase, you discover that you have been given more goods than you paid for.
- (b). They would never be missed, and their return involves the dismissal of the employee responsible. Would you return or retain them?

## No. 23. R.

- (a). You have made a true but damaging statement about an individual who is pronouncedly hostile to you:
- (b). Would you rather apologize to your enemy, or make a public defence before an unsympathetic audience?

## No. 25. R.

- (a). You wish to attend Science Courses which are available at either of two Colleges.
- (b). One offers good instruction with poor apparatus, and the other indifferent instruction and a fine laboratory; which would you choose to attend?

## No. 26.

- (a). You have invited out to dine a new acquaintance, whom you wish to impress favourably.
- (b). On finding you have forgotten your money, would you rather borrow from your companion, or risk the unpleasantness of an explanation to the management to whom you are personally unknown?



Examples of "Judgments" are shown in the following:

No. 16. R.

- (a). Longfellow immortalized many characters in his verse.
- (b). Which do you think the more popular, "Hiawatha" or "The Village Blacksmith"?

No. 21.

- (a). Pass in review the many inventions of Thomas Edison.
- (b). Would you consider Electric Lighting or the Phonograph the greater achievement?

In addition to the customary introspection the following estimations of degrees of confidence, weight and difficulty of decision were asked for.

(1). Grade your decision by the following scale, according to the degree of finality or confidence with which you would be prepared to act on it:

- (4). Perfectly confident.
- (3). Confident.
- (2). Some degree of confidence.
- (1). Not really decided.

(2). Grade the importance of the decision to you personally by the following scale, according to its gravity and weight.

- (4). Extremely momentous.
- (3). Serious.
- (2). Of some weight.
- (1). Unimportant.

(3). Grade the difficulty with which you arrived at the decision by the following scale.

- (4). Very difficult.
- (3). Difficult.
- (2). Of slight difficulty.
- (1). Easy.

Various forms of the demand to change the decision arrived at, of which the following are typical, were tried out in practice tests.

- (a). "Your decision is aesthetically wrong, try to bring your judgment into conformity with that of others."  
and,
- (b). "Now endeavor to make a reversal of your choice, and arrive at a contrary decision."

In the end it was thought best to keep this stimulus uniform throughout, rather than gauge the effects of varying forms of it, so that finally the latter form was adopted for all reversals.

After the attempt at reversal, and at the conclusion of the subject's introspection, the following estimate as to difficulty of reversal was asked for.

"Estimate the difficulty in reversing your decision by the scale below;"

- (4). Utterly impossible.
- (3). Barely possible.
- (2). Reasonably possible.
- (1). Easily possible.

The stop-watches used in the first series were fastened upright in convenient wooden stands, with the "heads" projecting slightly above the top so that they could be easily pressed. At the same time they could be conveniently read by the experimenter. By placing them adjacent to each other, No. 1 could be stopped and No. 2 started by the same movement of the hand. The average degree of error found by reading the differences in a series of twenty simultaneous startings and stoppings was found to be only .05", an infinitesimal amount when the greater inaccuracy of communicating the time of decision and the necessary stopping of the watch is taken into consideration.

The reagent was seated comfortably in an armchair, and, after being shown a sample of the problems, the following instructions were read to him: "Make your decision for good and sufficient reasons from one of the possibilities given. Work deliberately throughout and do not hurry over any part. Try to decide so that the situation appears as your own, and its contents a personal matter. At the same time do not labour over the decision, but let it be formulated as naturally as possible. When completed, pause to fixate your observations before communicating them to the experimenter." Then an example with a cover page was handed to him, and when he turned the cover, a stop-watch was started. When the first part had been read through and the situation grasped, the subject turned to the second part; then the first watch was stopped and a second one started simultaneously. When the decision was arrived at, the subject signalled to the experimenter, and the watch was stopped. This was found to

be more convenient than asking the subject to attend to this matter, for it in no way seemed to contribute to the acting out of the decision as was suggested in the work of Michotte and Prüm. In some cases in the trials, the tendency was to forget to press the spring in the absorption of the choice. After a short trial series had been worked through by the subject, the regular series was attempted. When the decision had been arrived at and the introspection written out, the subject was instructed to evaluate his choice in the terms previously indicated.

In the second series, where choice of odors for smelling by the subject was made, the series consisted of eight liquids coloured for their identification by him, and numbered for the same purpose for taking records by the experimenter.

They consisted of :

<i>Number</i>	<i>Solution</i>	<i>Colour</i>
(1).	Asafoetida Weak $\frac{1}{4}$ strength.	Light Brown.
(2).	Valerian Extract $\frac{1}{4}$ strength.	Carmine.
(3).	Eucalyptus.	Clear White.
(4).	Vanilla.	Violet.
(5).	Oil of Cloves.	Green.
(6).	Valerian Extract Full strength.	Dark Brown.
(7).	Lysol.	Medium Brown.
(8).	Asafoetida. strong.	Dark Chocolate.

They were presented in pairs, by means of an exposure apparatus consisting of a small cloth screen, which was suspended in front by means of an electrical magnet at the top of a stand. A stop-watch was so arranged in a fixed support that when a short projecting lever arm pressed down the watch to start it, contact was also broken in a circuit, allowing the curtain to fall and exposing the bottles. The subject was required both to start the process and stop the watch on arriving at a decision. Before actually commencing work with the series, the subject was practised a number of times in starting and stopping the apparatus, and when proficient, proceeded with the regular series. Before actual experience taught him the olfactory values of the contents of each bottle, he had only the distinguishing colour to guide his choice. Later, the colour became associated with the actual odors, and



thus served as an indicator of the olfactory quality, and hence, for him, of the real value. After a decision had been arrived at, the subject was required to give an introspective account of the process, and to evaluate his reactions as in the former series, as to degree of confidence in his decision and difficulty of arriving at the choice, the degree of gravity being omitted in view of the fact that all choices in the series were of the same type throughout.

Requests to reconsider decisions were then made by handing to the subject a typewritten slip bearing the legend previously used: "Now endeavour to reconsider your decision and make a reversal of your choice."

The time for this second reaction was noted as before, and after introspection, evaluation for "Confidence and Difficulty" was required, as in the case of the original decisions. The odors were presented once in turn in each of six rounds. By excluding each as it was smelled by the subject, and by a careful process of juxtaposition of the remainder this was rendered possible. Two rounds were worked through in each sitting, the order of succession being different each time. Thus three sittings were completed by each subject, at intervals of from three to four days between each. The subject was in no way informed of these arrangements, and the irregular order of presentation tended to keep him more or less in ignorance of the procedure. At the end of the second sitting the subject was asked to arrange the bottles in order of merit, in so far as his experience guided him, and this procedure was again required at the end of the third sitting. When this last arrangement was concluded, the subject was again required to arrange them by means of a rating of smell values without regard to any other values. For this purpose any direct comparisons that he needed to make in order to be quite certain of values were actually made by him. This final scale was used as a basis for checking up the two previous scales.

While the first series attempted an investigation of situations from practical life which, however, were impossible of actual realization in the laboratory, the second series seemed to offer a ready method of comparison with it. On the one hand was a process arising from a hypothetical situation and, on the other, a practical problem (even though trivial) realizing itself in action. In this way a control set of one against

the other was secured. By the method of following up the choice with a demand for reversal, not only was the period of choice subjected to examination, but the after period and the nature of the mental set with its possibilities of modifiability were also brought under investigation.

### III. QUALITATIVE ANALYSIS

Before attempting a detailed analysis of the various developments of the process and their emergence into defined types, an illustration that presents the processes as a whole provides not only a sample of the method of procedure, but also shows the need for detailed discussions of the progressive stages. In the succeeding paragraphs there are given two typical samples of introspective results—one from each series—together with the time and the subject's ratings as to confidence, difficulty, and so forth. S. I. No. 10.

(a). You are unprepared for a change in the weather, which has suddenly turned cold and rainy.

(a). Time, 9.2".

"It was represented by a street scene; a point on the left side signified my location. My view was to the right at that point. All that was present was a downpour,—definite kinaesthetic and visual streaks from above down on the visual field. There was also present a second part to which I turned my attention, i.e. 'unprepared.' I cannot say how I thought of it, but I seemed to feel it as a 'thinness or lightness' with a general awareness of my skin. This feeling was not a temperature image, but like the experience in Spring, when one takes off heavy winter garments,—hence there was a motor feeling of lightness."

(b). Your way of getting home is by riding on top of a "bus", and thus risking pneumonia, or in an overcrowded subway train with an influenza epidemic at its height: which would you take?

(b). Time: 29.6".

"First came a faint visual picture of a bus pointed towards me, with the awareness that it meant coming 'up town' as opposed to 'down town.' This was followed by a vague outline of the top of a bus and the bodily presence of myself in the back right hand corner. The most conspicuous parts of this image were my shoulders and especially my chest. That disappeared, and the second alternative took form as a small area of a subway crowd, not larger than I could reach out and touch with my hand. I was the centre of this area. There



was a definite kinaesthetic 'suggestion' of restiveness and a visual element with vague patches, meaning, 'people in close contact.' This was associated with general bodily discomfort and a feeling of closeness, with a shrinking from them. The word 'influenza' was followed by a third image, with the consciousness, that this was closely related to the second rather than the first. Human faces appeared wearing handkerchiefs or influenza masks; there was a definite kinaesthetic sense of direction to the north, meaning, P..... Then, in a curious way, I reviewed the whole episode of the influenza epidemic at P....., and the wearing of the masks, but present with this was a definite notion of calm and security. This was in contrast with the emotional tone of the first image (chest and shoulders prominent), which was accompanied by mild fear and avoidance. I then reverted to the words on the page. Having been represented the images passed out, and I fixated the phrase, 'risking pneumonia', that was accompanied by the same mild emotion of fear. Then I reverted to the image of the white masks, and this constituted the decision. The feeling of the act of coming back to this was recisive. Underneath all this was something not represented in imagery. Verbally it was, "You once had pneumonia and it might be dangerous for you, whereas you went through the 'flu' epidemic without any harm. With this was the attitude, 'I must take this reference to my past experience as crucial and determining this present decision.' I cannot say how articulate this was, but it was decidedly present."

Confidence—3 or confident.

Weight—4 or extremely momentous.

Difficulty—2 or of slight difficulty.

(c). Now endeavour to make a reversal of your choice and arrive at a contrary decision.

(c). Time: 38.6".

"The Aufgabe was accepted with great reluctance and an undertone of annoyance. All that I seemed able to recall here was, what seemed to be auditory in form, the word 'pneumonia', accompanied by a definite feeling of dissent. There was a distinct sensation of sinking back on my decision when I heard 'pneumonia,' but when I considered the other alternative there was a rising towards it: then the audition of 'pneumonia' seemed to make me sink back again. The final

decision appeared to be that of a 'falling back,' probably a relaxation of muscular set. There was, throughout, a distinct feeling of effort accompanied by the meaning, "I must honestly try to do this," which involved the 'rising up' towards the 'pneumonia.' It appears, judging from the protracted time that it seemed to take, that more must have gone on here, but I can't say what it was." (The experimenter then asked the subject to compare (b) and (c) in regard to degree of effort involved and amount of clearness). "Part (b) involved no effort but (c) required a marked degree of effort. Part (b) was very much more vivid than (c)."

Possibility of Reversal—4, Impossible to reverse.

### Series II

S. IV. Round 3. Odors presented were, on the right, No. 4. Oil of cloves, colour green, and left, No. 5. Vanilla extract, colour purple.

(a). Fore period. "An attitude of expectancy and curiosity was felt, accompanied by marked tension of the throat and vocal chords."

(b). Main Period. Time 3.2".

"As the curtain fell there was a marked feeling of tension and suspension of breathing. On seeing the bottles the green one was noticed first: I looked back and forth for some time and remembered I liked both, but almost at once decided on the greenish one as having a much more pleasing odor. With this came a faint, olfactory image of 'wintergreen' and this seemed to play an important part in the decision. The one chosen seemed much more attractive than (5), which was regarded with indifference rather than with repugnance. With the decision there was a definite kinaesthetic act of choosing. There was a kinaesthetic image as of going away from the purple towards the green. The decision was accompanied by a sudden liking for the one chosen which made it stand out very clearly. The final act was accompanied by marked satisfaction. There was, after this, a definite kinaesthetic 'set' as of being about to reach out towards the one chosen and then smell it."

Confidence—4, perfectly confident.

Difficulty—2, of slight difficulty.

(c). Now endeavor to make a reversal of your choice and arrive at a contrary decision.

(c). Reversal Period. Time: 6.2".

"I merely glanced at the paper to reassure myself of the usual instructions, at the same time with pronounced feelings of irritation. Then I looked again at the exposure stand. I paid particular attention to the purple bottle and tried to call up an olfactory image of it so that I might put it on a level with the green. It refused to come, but the image of the green bottle returned, though not so vividly as before. It was vague but I endeavoured to cling to it, for it seemed that, if I lost it, I would have to accept the purple, which I really did not want to do, for having reached the decision the first time, I did not like changing again. The effort of attempted recall was very strong, but the 'purple' odor refused to come, and I felt that the image of the green was weakening, so I decided to reaffirm my former decision. The satisfaction here was not marked. Not only was I irritated by the demand to reverse, but there was also a marked feeling of annoyance at not being able to recall the olfactory image of the purple.

Confidence—4, perfectly confident.

Difficulty—1.5, between easy and of slight difficulty.

From a perusal of these examples it is evident that the conventional stages of fore period, main period, and the after, or reversal period are obvious and convenient divisions for discussion, and contain within themselves well defined processes, divisible, where necessary, into sub-processes.

#### (a). *The Fore Period*

In the first series the fore period takes on a primary importance, the task or Aufgabe demands not only a general mental set as a preparation for (b) the main period, but the calling up of a situation that in large measure determines the factors of choice. Thus while the outlines of the actual situation are prescribed, the details are filled in subjectively. It is for this reason then, that in the first series, the same situation ranges from "trifling" to "serious," or of "some weight" to "momentous" in the same examples, according to the estimation of different subjects. That which appears simple and trifling to one individual, seems to another to assume grave importance. The fact coincides exactly with similar experiences at the different stages of one's life, that which is, at one stage, of serious moment, appearing at another an insignificant detail. The degree of seriousness depends, in the main, as to how far



such an occasion has previously entered into the subject's experience, e.g. whether entirely novel, or similar in type to some previous experience frequently met with, and secondarily, on the disposition and mood of the subject.

The adoption of the Aufgabe was at first conscious, but afterwards became less focal in character; there was a definite striving to formulate the situation so as to meet its demands, which was general to all subjects.

S. I. No. 4.

"I read over the sentence twice and noticed that my attention was called to the word 'limited.' I turned back, having a kinaesthetic sensation of the eyes in doing it. For a moment I dwelt on the word and its meaning. In thought arose the question, 'How limited?' After fixing this and raising the question I decided to go ahead. I had a vague feeling of dissatisfaction and dismay at the indefiniteness of the situation as I was conceiving it, but I decided to go ahead."

S. II. No. 11.

"On reading the facts of the situation there came to me a visual image of a man's dress suit, but I corrected it so that it became an image of a woman's evening dress. The thought came that it would be more appropriate to decide from this point, and then I had a visual image of an evening dress of my own."

S. III. No. 2.

"When I read the words the situation left me cold and unmoved. My attitude was matter of fact, and I accepted the situation only in thought. There was a slight feeling of dismay and anxiety that I had not managed here to pin down a real situation, so I began to cast round for a something more real and definite. Then I commenced to imagine myself at the corner of 116th. street and Amsterdam Ave., I had a visual image etc."

S. IV. No. 5.

"With the reading of (a) came a moderately vivid kinaesthetic image of tiredness in eyes and head. (I feel tired sometimes in the eyes.) There was some effect on the breathing which seemed to slow down. A relaxed feeling more pleasant than unpleasant stole over me, and I visualized myself in my room in a Morris chair. Suddenly I realized that this was not completing the situation, so that my attention was directed to the outside of my room, and I faintly visualized a storm with thun-

der and lightning. There was an idea of the rain also, that must have been present as a very faint auditory image."

S. V. No. 3.

(a). You are attending a match in which your side is winning.

"Had a feeling of expectation before turning the paper; on doing so read through the whole with internal speech, the word 'match' seeming to stand out. Immediately there flashed into my mind the visual image of a basket ball game that I had recently attended and next of a tennis match. The latter was very vivid and I waited awhile recalling visually some of the strokes that were played. But the thought came 'I must not waste time but fix the situation', so I determined to keep to the tennis match. I thought, in meaning, "This will do nicely, so I'll make it the tennis match.'"

Later as the process became more mechanized and the reagent more confident, the tendency was to seize the main points in a given situation without delay, and if the situation did not readily realize itself, to pass on and rely on further details such as would be supplied in (b), to formulate it more definitely.

S. II. No. 7.

"I could think of no particular experience that would help me to a realization of the situation, but it formulated itself in thought without the interposition of any perceptible images. Having pondered it awhile I turned the page to see what the alternatives were."

S. IV. No. 25.

"The term 'Science Courses' was not clear, I thought of them merely without any image. 'Two Colleges' also came by thought process alone. When I had related the two ideas, I had a thought that it would be best to pass on to the next part, and then I would be able to get the whole situation clear in my mind, so I passed on to (b)."

S. VII. No. 23.

"I had a difficulty in forming a definite picture of the situation suggested in (a) since it seemed so complicated. I thought over people that I knew—mainly by means of visual images—but I could not think of one to whom I might apply it; I therefore knew I would have to invent one. I thought of a friend of mine whose father was a physician, but I waited until I had read (b) to make the whole setting more concrete."

## S. VIII. No. 17.

"First I thought vaguely of our home in the country with a vague visual image of part of the inside of the house,—the living room and library, where I proposed to hang the pictures. I had a curiosity to see the pictures themselves, so I turned to (b) knowing that I would better be able to decide on the place to hang them when I had seen them."

This is, perhaps, the more correct psychological procedure for, in a practical situation, the tendency is to formulate a decision along with the focussing of the various factors of a situation, rather than to perform the two acts separately, cognition being a concomitant of, not a preceding act to, volition. At the same time, the realization even in bare thought of such a situation as the fore period (a) describes, tends to create a background for the situation, and thus readily permits a focalization of the necessary factors.

Frequently at first, with the formulation of a situation that met the demands of the task, there followed a slight pause marked by relaxation of effort, a relief from tension and well marked feelings of satisfaction with the achievement. It corresponded to a halt in the work to "note progress." This was general to all subjects in the early stages, but became less prominent with the mechanization of the process.

## S. I. No. 6.

"I felt I had a good situation clearly in mind and felt pleased, and the feeling of strain lessened considerably for the moment."

## S. III. No. 8.

"Was glad I had made such a suitable response and I relaxed to some extent."

## S. II. No. 18.

"As I turned the page, the thought came that I had fixed the situation so that I would be ready to meet the alternatives."

## S. IV. No. 10.

"Thought this idea met the situation very well and was very well satisfied; at the same time the tenseness in the frontal region temporarily disappeared."

## S. VI. No. 3.

"I knew that I had fixed it as well as I could manage."

## S. VII. No. 5.

"I thought, 'this will be just the thing.' As soon as I had realized the situation the strain lessened considerably."



In some other early cases before mechanization of the process, in place of the pause of satisfaction, the tendency was towards an anticipatory state; expectancy, curiosity, and even apprehension entering in. In regard to the latter component, a direct shift of tension about the region of the diaphragm accompanied it. The other conditions are corroborated both by Michotte and Prüm (13) and Wheeler (24), and by results obtained in the second series of the present experiment.

S. I. No. 17.

"Thought perhaps they would be like the two landscapes of the previous experience. With this a sense of satisfaction came over me, and I felt the muscles of my face relax as if I was smiling."

S. IV. No. 6.

"Having done as well as I could manage, I was impatient to turn to the next sheet to observe what was there."

S. VI. No. 9.

"Fixated the situation and was anxious to see what was coming. Wondered what the alternatives of this situation might be, and hoped they would be pleasant. Did not like the situation at all, and wondered if the possible alternatives would be very unpleasant. There was no relaxation of the strain in throat and forehead, and there was a sinking feeling as of dread about the diaphragm, just as if something I did not like was about to happen."

In two cases subjects reported having arrived at a solution to the situation before even the main period was reached, the state of anticipation having realized itself without delay.

S. III. No. 9.

"A visual image (of my friend) quite distinct in detail, came to me, and I thought of what might be preferred. Immediately some pictures that I had seen in the bookstore came to me and I thought verbally, 'Those will be the very thing.' Then I remembered I could only choose from the pictures submitted, and hastily turned to (b)."

S. VIII. No. 6.

"Saw myself at 42nd Street Station just about to enter. Just then I imagined myself going up to the desk, and looking down saw my pocket book was missing. I felt dismayed, and stood a little while undecided as to what I was to do. I had a kinaesthetic image of tiredness in my legs and head, particularly about the eyes. Suddenly I thought of a taxi, and the

thought came with it, 'I'll pay "him" (the driver) when I get home.' I felt and saw myself walking up to Broadway and calling up a 'Black and White' taxi, since it would be cheaper; and next I found myself at home going upstairs to get the money to pay the driver.

Then I turned to (b) and, on reading it, the image of myself and the taxi 'crashed out' of my mind; it left me just tired, helpless, and, for what seemed a long time, I was unable to get a new 'set' for the suggested alternatives. They came very slowly and with difficulty."

These cases corroborate what has already been suggested as to the inseparableness of the statement of the problem from the consideration of the alternatives for its solution; the two are as involved in their operation as cause and effect, like them being separable only in regard to their theoretical consideration.

*(b). The Main Period.*

The main period comprises two well defined divisions, distinct in their processes and their products; the former may well be termed associative, the direct product of the Aufgabe, while the nature of the latter is characteristic of, and peculiar to, decision alone.

The associative side may be again divided into realization of the alternatives, which implies also a degree of evaluation, according to their feeling tone, and the continuation of this process of evaluation into one of comparison. There is found no hard and fast line of demarcation between these two minor processes, for they merge imperceptibly. The stages are distinguishable, however, if not separable. The two processes are essentially intellectual, and, to use Wundt's term, "apperceptive" in their character, for there is a definite attempt in all cases to first link up the preceding period (a) with the factors suggested in the present, (b), and to evaluate them in accordance with past experience.

S. I. No. 19.

(b). "The first thought that occurred to me was a specific occasion of having previously been given goods I had not paid for, and with it a vague visual image of the place where this occurred. I read the next clause,—'their return involves'—here came a vague visual representation of the store and the path leading to it, but the 'store' end developed. There was a

vague rectangular representation of the store with two people in it. One man was near the door while another was further back,—employee and employer. About the middle of the sentence I made a tentative decision, somewhere about the word 'return.' 'Certainly I will,' came the thought. Then a curious state of mind came up. It involved a discussion of the whole of McDougall's 'Social Psychology,' and its consideration of the development of moral ideas. I had a diagrammatic representation of the development of the topic, which came as a kind of background in connection with this. It signified, 'Here is a case of what developed in this diagram, but something is certainly withheld here.' This was not articulate but present as a thought. As I went on, 'involves the dismissal of the employee,' the schema was replaced by an image of the store again. The focal thing now was the man near the door, and for him I had a distinct feeling of pity. With that came a feeling of negation and reluctance on my part."

S. II. No. 10.

(b). "At first a visual image of a green 'bus moving along in my direction come fairly clear to me. The upper part was very distinctly outlined, though the lower was somewhat vague. This was accompanied by a cold image and I felt myself shiver. Next appeared a visual image of the 'subway' as it appears at 14th Street. Crowds of people seemed to be surging past me. They were coming from above,—that is to say, the top of the stairs. The 'bus image alternated with that of the subway. Then came a digression in thought concerning 'flu' and its causes. The thought came that the average person will tell the cause with certainty, while the doctor hesitates. I then turned to the alternatives again, and felt my eyes moving to and fro, now fixed on the bus and now on the subway entrance. This occurred at least four times. With the bus was a disagreeable feeling of cold accompanied by shivering; with the subway came smell images of the crowds and of the subway itself; these were also disagreeable."

S. III. No. 11.

(b). "To the question—'Would I prefer plain or evening dress?'—there was an immediate feeling of stress and strain evoked by the situation which was very unpleasant; it was most pronounced in the thorax, and I noted my breathing was suspended. There was a decided feeling of annoyance and



irritation. Immediately I felt myself attempting to escape by looking for a chance of evading the problem by thinking, 'I'll find out what others wear.' This was, however, evidently not allowed by the conditions; I knew that I had to go without finding out, so I dismissed the former idea. Then I re-read the question, and the thing that stood out prominently was the strong dread of being conspicuous. A feeling of hurt pride and dignity that was most painful and annoying, increased whenever I thought of this phase, so I tried to avoid it. I pictured myself in plain dress with others in evening clothes, this was soothing to some extent, and then I reversed conditions and was not at all satisfied, for this seemed to make me more conspicuous and hence the annoyance increased. I tussled with the situation to and fro for some time before deciding."

S. IV. No. 10.

(b). (Standing at Columbus Circle),—"I had a visual image of a bus coming towards me in the rain. I had a sensory image of being cold and chilly, and felt myself shiver. The thought of pneumonia brought on a feeling of fear located about the diaphragm. The thought came non-verbally, I would have liked the bus very well in fine weather, but now it is too rainy. Then I turned my attention to the subway. This felt warm—here I had a slight sensation of warmth, and besides, I knew I could not get wet. The rain now seemed present outside as an auditory image. Then the condition of overcrowdedness came and I felt strong objections to going home by subway. To this was added the risk of influenza, but the idea came that this might be prevented if precautions were taken. I then considered the two situations, pneumonia vs. influenza, for some time."

S. V. No. 2.

(b). "With the reading of (b) came first, a definite visual image of a mirror. This was transformed into smooth ice, and then ice with skate marks on it. This series had a pleasurable feeling. The 'falling' was only present as an awareness, unaccompanied by image of any sort or any feeling tone. 'Slush' brought up a kinaesthetic image of stepping into half-melted snow, and I thought of a situation where I had to get a car or miss an important appointment. This had occurred to me yesterday, and a fleeting image of losing my car and the dis-

agreeable feeling of annoyance and dismay that followed, surged over me again, and I turned with relief to consider the ice problem."

S. VI. No. 13.

(b). "When I read (b), I thought neither alternative fitted the situation; they came only as thoughts and not as specific cases; at the same time there was a strong aversion to either. Then the idea came in meaning, 'Well, perhaps I can give the story such a generalized bearing that the individual would not be aware of any personal reference,' but I knew this was an evasion of the problem. Then I tried to imagine myself in such a situation in which this would be impossible, so that I would have to go on with the original story or stop. I tried to think of people I knew, with a special reference to some points they were 'touchy' on. There was a strong sense of strain, mainly located in the forehead and frontal region of the head, while breathing was slow and suspended. I dismissed one person I knew since she was merely cynical. Next I thought of a neurasthenic that might be satisfactory, and I had then to think of a case in which the alternatives could come up. I found it and the tension was considerably relieved. As I went on with the story I knew how unhappy this would make her, and I felt unhappy in sympathy with her discomfort. She was in the centre and I saw her clearly with the group round, the latter being rather indistinct. The situation was particularly vivid, and I felt I had arrived at the required situation. It was exceedingly difficult, for I felt chagrin at what everyone would think if I broke off in the middle, for the very breaking off would imply some personal reference within the group; as a result, the woman would be self-conscious and the whole group would feel uncomfortable. When I turned to the other side, I felt that if I went on, the fact that the woman was neurasthenic would make her feel the difficulty very keenly. The only possibility I could think of was to turn the specific reference into a general one, so that she would not think it aimed at herself, and I considered this for some time. The whole occurrence was accompanied by considerable strain and distress of mind that was very unpleasant."

S. VII. No. 10.

(b). "First a visual image of myself riding on top of the bus. The feeling was pleasurable, and the notion came that I

was warm-blooded and I could stand the cold. Then the idea of pneumonia came and affected me considerably, and the former pleasure faded. In the background was a reference to the fact that I had already had bronchial pneumonia and this coloured the situation very strongly. Then I turned to the subway. The thought came non-verbally 'I have avoided pneumonia and I do not want it now.' There was also a background of dislike towards this means of getting home. I had a partly visual and partly kinaesthetic image of myself being wafted along in the crowd, and entering the car at the centre door. Then for a while all the imagery faded and my mind appeared to be a blank. I made a strong effort but could do nothing till I re-read the words, and recalled the imagery with great difficulty. It returned as I read the words, and was accompanied with a markedly disagreeable feeling tone. The effort made itself felt in a strong sense of strain in the frontal region of the head, less marked in the chest, and with suspended breathing."

S. VIII. No. 17.

(b). "On turning to the prints (b), I observed the one on the left first, since the brightness of its colours seemed to attract my attention. Then I felt my eyes turned to the other which was very much duller. Pronounced feelings of pleasure came with the observation of both of them. The one on the left stood out more vividly and forced itself on the attention, but the liking from the first was stronger towards the other on the right. The definite thought came that I liked its dulness and general unobtrusiveness, together with its deeper perspective. I had not yet decided, but turned to the one on the left to attend to it again, and its brightness returned with renewed pleasure, but I felt the Corot (right) was the more attractive with its quiet conservative tone."

For introspective examples illustrative of this stage of the second series, the reader is referred to the discussion on "motivation" (d), where the matter is treated in detail.

If, now, this part of the decision process be compared with the corresponding side of the judgments, there is found, throughout all four cases given to each subject, two main differences: on the one hand, in decision, the subject himself is always imaged, or else thought of in connection with and forming an integral part of the general situation; in the case of judgments all reference to the subject may be omitted, and



he may remain quite out of the picture. If the images are very vivid in the former case, the subject really lives through the tentative experience, while in the latter, he remains an abstract entity, an arbiter, and by no means a necessary part of his mental pictures; the content being associative and perhaps of a non-personal nature. This essential difference of the subject as content, or abstracted from content, is accompanied by differences in emotional and feeling tone which follow as a necessary consequence of the preceding. Thus situations that involve decision sometimes imply emotions of dread, dismay, exhilaration, and a more or less marked feeling tone of agreeableness or disagreeableness with accompanying stresses and strains. In the case of judgments, the reagent may remain calm and unimpassioned, with no marked stress or strain beyond the tension in the frontal region due to fixation of attention, or, in the case of S. III., of laryngeal tension due largely to the fact that her attention processes were verbal motor. A comparison of the succeeding specimens of typical reports will show the difference, even here, between the two types.

S. I. No. 22.

(b). "Here everything seemed to be based entirely in terms of meanings and thoughts. If verbal elements were present my attention was not on them. My thoughts were to this effect: The primacy of Smith's attempt was just an accident of time. There seemed to be a vague diagram at that point, a line running from the distance towards me. A point or mark on that line designated Smith's enterprise. The meaning was, 'If it did not come then, it must have come later along the line.' Turning to Washington, this was represented by a vague image of his portrait. In addition, there was an awareness that it seemed to be in the air and to stand over something that I conceived beneath it, but this something was not pictured. The meaning accompanying it was that of 'stability,'—there was no movement to it. That meant, for me, 'This work is of a qualitative kind'."

S. II. No. 21.

(b). "When I turned to (b), I felt very satisfied and pleased with myself that I had thought of the same two things as the experimenter. But the thought came that this self-satisfaction was not the point, but rather that I should attend to the consideration as to which invention was the greater. Then the question also came as a thought,—'Greater in what—difficulty

of achievement, or greater as regards human welfare. I determined to decide on it from the point of human welfare, and there came at once a visual image of electric light pictures and of electric lights themselves. The phonograph side was also visually represented by disc records, and did not have much of a chance to come clearly."

S. III. No. 22.

(b). "I took up the question in chronological order. The chief process was ideational, with faint references to vague visual images. Smith was thought of as a positive force in the development of the nation. His contribution did not appear an essential one, he but hastened the work another Smith would have done. Throughout this thought came a vague visual image of Smith in Elizabethan dress among others who seemed to be present in the picture. This was probably a memory image of an old school book. With the thought of Washington came a vague reference to his portrait, which is common in most public institutions. In thought came the idea, 'The work of Washington required genius and personality. Without him, it is doubtful whether the thirteen isolated colonies could have been welded into a whole.'"

S. V. No. 21.

(b). "At first came only a realization of the situation in thought, and the question came up as to whether I must decide from the point of the actual achievement in invention, or as to its after utility. Then the idea appeared in meaning, that others also worked on and improved both the phonograph and electric lighting systems. With this thought came a clear image of the advertisement picture for Mazda lamps. The thought came, 'Well, other electric light companies may provide as good lighting as Edison's bulbs, and the idea followed that a Sonora or Aeolian machine was as good as Edison's.' Here came a distinct visual image of a gramophone. I did not know of the history, or Edison's part in the invention however. Just then I returned to the first question as to which point was to be considered in order to come to a decision. Finally I decided, what afterwards seemed obvious, that, as I did not know the history, I could only decide on the ground of utility. I then began to compare them in this respect."

(b). "Turning to (b) my curiosity was satisfied. I saw in visual imagery Hiawatha and The Village Blacksmith as per-

sonalities together for a fleeting moment, then I was ready to go on. Considering Hiawatha first, there came to me with an auditory image the rhythm of its verse, but I could not catch any verbal imagery. This was followed by a visual image of a body of water, (evidently the reference was 'Gitche Gumees—Big Sea Water'). Turning then to The Village Blacksmith, a distinct visual image came up; the man was standing near his anvil within a vague setting of a blacksmith's shop. I pondered the question in thought, and the idea came, that, if Hiawatha was more popular with children, The Village Blacksmith was more popular with their elders. I rejected the idea that on this account either would be more popular, and tried to consider who would read them. Then it seemed that though children read and memorise Hiawatha, it soon 'gets away from them.' It is an emotional attitude rather than an intellectual one; while, on the other hand, The Village Blacksmith has more content, a stronger appeal to maturer minds, and sticks longer. I weighed these points some time before making up my mind."

S. VIII. No. 28.

(b). "On reading (b) an element of surprise entered in. I had not thought of religion in (a) but of home and school. For a time, the originals passed out and I had to arrange a new mental set to make the comparison. The other faded quickly and was succeeded by a blank period, accompanied by a feeling of disagreeableness and tension in the forehead. Then came into my mind the influence of the Roman Catholic church, typified by a mental image of a priest in clerical dress. This stood for the far-reaching influence of the Roman Catholic church. Next the bare thought seemed to come, 'I do not know enough about religion to settle this.' Outside the Roman Catholic church, perhaps the school may have more influence. Then came up a visual image of a schoolroom, different parts of it looming up in rapid succession, as if I were viewing different aspects of it. I, however, was not present."

In these differences, already apparent, must be found the germ, developing later in final decision and judgment,—of the whole process of self-consciousness. The subject in decision here begins to realize himself as part of the alternative situations, which are finally to develop into real activities, and which he now acts through mentally before committing him-



self to either. In the case of judgment, the reagent remains a passive onlooker throughout, taking no part except at the end to declare the result of the final count.

The period of decision appears to yield three distinct types, each with a characteristic reaction, and possessing some points of identity with those noted by Michotte and Prüm (13). They have been designated here from their more important characteristics, as the Preference, the Conflict and Indifference types: in addition there is also the fourth class, that of the Judgments, which are not choices in any sense of the term, and they, too, require a separate treatment. The first and second appear to correspond to Michotte and Prüm's "Voluntary Decisions", and the third to their non-voluntary or automatic type. These types are also described by Wheeler (24), though not specifically designated by him, and may be found exemplified in a study of the examples of introspections supplied by him. Still, in following up the effect of the after period, it is necessary, for the purpose of the present investigation, to note and designate them.

The *Preference* type appears as the predominant form in which the subject arrives at a voluntary choice. It follows from a clear concept of values of the alternatives and a possibility of their comparison, with a final summing up of the motives. They may be either "negative" according to whether both alternatives are more or less disagreeable to the subject, when one is rejected and the other accepted, or positive, when one is more markedly pleasant than the other, when the alternative attended by the more pleasing motives is eagerly accepted. The degree of certainty with which these motives are apprehended, and the magnitude of the difference between them, are important factors in hastening or delaying the actual decision, and also affect the degree of confidence in the final act of choice. In some cases, there is an initial impulse to one or other alternative—less rarely to both in this type—with only sufficient delay to focus attention on the other alternative to note it, and then to return to the original. In other cases both alternatives were fully reviewed and weighed before final commitment. The former occasions were those where the situations were familiar to the subject by analogy with recent occurrences, in the latter, the situation had to be more laboriously constructed. Nevertheless, the

final act of decision was characteristic for all. There was a tendency for the attention to fixate the chosen alternative, apparent either through an increased vividness of the image or, where the object was concrete, by a concentration of the gaze upon it. In the second series, that of odors, this fixation was an obvious factor of the subject's behaviour, and the experimenter could always tell which odor was chosen by the subject, by noting the direction of the final fixation of the eyes.

Accompanying the attention process were various kin-aesthetic images. In the case of the first series the subject sometimes described the effect as a positive attraction or pulling power of the one image representing an alternative, or a positive repulsion from the other, according to whether the choice was positive or negative in character. Thus there was a tendency "to go forward to," "to assimilate," "to coalesce with," the image, "to identify it with the self," or, on the other hand, "to reject," "to put out of the mind," "to get away from," "to shut out," the more repellent image. Corresponding to this in the second series was a tendency "to regard with pleasure," "to reach out and seize," "to take possession of," or, on the other hand to "reject," "to violently repel," "to brush aside," or "to push away" the rejected bottle. Following upon this was a more or less marked degree of satisfaction with the choice, a relief from tension, e.g. bodily strains, suspense of breathing and tension in the frontal region or throat. Such accompaniments were the less marked where the choice was easy, or of the character of following an initial impulse, but invariably accompanied the more prolonged and strenuous type. Preceding the act and accompanying the bodily tension, the subject was often bowed forward with the head bent, but immediately after the formulation of the decision there came a straightening up and settling back in the chair, and afterwards the release of tension and the resumption of normal breathing.

Finally, synchronising with the decision point, there was a tendency to justify the alternative chosen, by specific reasons, which often arrived with a "rush," thus enhancing the satisfaction of the choice. This process may well be classed as one of "rationalization," in the commonly accepted psychological sense, i.e. a process of justification for the choice after its completion to satisfy oneself rather than of affording

logical presentation of reasons to influence the process beforehand. Where the choice was one of positive acceptance, the reasons were favourable to the object chosen, where the more marked action was one of rejection, then the tendency was to invest the rejection with some negative reasons derogatory to the rejected alternative.

S. I. No. 5.

(b). "There came a visual recollection of a little table and I felt myself sitting down to that table with a notion of, 'That's what I very often do.' I had a feeling of certainty with respect to the choice. There was a kinaesthetic reference to myself as seating myself, as well as the visual representation. There was no real deliberation as I did not consider the other alternative. The only 'other' elements present were a sense of ease and lack of effort, and a certainty with regard to my choice that was agreeable."

S. I. No. 11.

(b). "During these alterations I was aware of the notion that, on the one hand, I should have social standards on my side, on the other, I should feel I had made a mistake. The decision came as an eye movement, the fixation of the words 'evening dress' on the page. I was aware of a definite inclination towards the left, kinaesthetic in character. I had a definite feeling of control of the situation, a dismissal of the one alternative and acceptance of the other linked to the visual shift. The decision seemed as real, on the whole, as a real situation would be, and appeared as a genuine decision. The process was attended by tension in chest and front of the head, but this disappeared when I had chosen. I felt myself settling back in my chair after the stooping position maintained during the process of choosing."

S. II. No. 8.

(b). "There was a disagreeable feeling, and an attitude of disapproval in relation to the person hunched over the table and eating as hard as he could. The thought accompanied it, that it was not ideal or desirable to eat constantly by oneself. This decided me at once, and the image of this person hastily disappeared, leaving the idea, "Well, I've decided not to do that." There was very little tension, but between the attempt to decide and the actual decision itself, there was a suspension of breath and a moving of the eyes from side to side."



## S. III. No. 23.

(b) "The first reaction was one of indignation at the audacity of the question. There came in thought, 'The statement was true and the man an enemy, fancy apologizing for the truth!' This would be compromising a sense of right and wrong and would be cowardly. If I made the statement I would defend it and even take the chance of 'mobbing.' Here came a vague mental picture of a seething mass of humanity surging about me. But the idea came, 'This is too precipitate, so take hold of yourself and consider the situation.' I felt I had to make a distinct effort, located in chest and arms, as if I had to get up out of my chair to do it. I had to think over apologising. Immediately I thought of apologising by imagining myself before an audience, the strong feeling of dislike surged up and I felt I had to discard it utterly. Again the idea that the statement was true flashed into my mind, and I thought verbally, 'Why, the other side haven't a leg to stand on.' Immediately I experienced a kinaesthetic image as if I had launched myself at the decision like a football tackler after being held back. The tenseness disappeared at once and I felt satisfied. After the decision came, I felt an emotional glow of righteous indignation and expectancy, as if ready to meet the antagonist at once and get the affair over. My fists were clenched and my jaws were set in preparation. There was not the slightest doubt in my mind from the first as to how I would act. The only delay was to consider fairly the other possibility. This was the most tense and unpleasant period of the whole experience."

## S. III. No. 15.

(b). "For a moment I felt no particular preference but, as I studied them, I felt a tendency to reject the right hand picture. Then there began to come up reasons for the left hand picture as against the other. The idea came, "This is a real Dutch setting; the other is nondescript." The setting of the one was indescribably superior to the other. The one was bright and the other grey-looking. The decision came with a rush and a marked degree of finality. 'This is mine,' I thought, and ceased to regard the other completely; it was as if non-existent for me. The whole experience was accompanied by a minimum of strain and the whole process was even and 'gentle' throughout. At the same time I felt quite certain of my decision which was final and complete."

## S. IV. No. 4.

(b). "Had a visual image of myself hatless and at the same time with old worn shoes, which passed into fleeting image of myself with influenza. The other parts were indistinct. Then the image of the shoes grew brighter, and a temperature and touch image as of myself paddling through water came vividly into my mind. The hat had completely faded out but the reason came in thought, 'My old hat is still wearable, but not these shoes; I'll have the shoes.' The tension in the forehead disappeared at once, and a pleasant feeling of satisfaction came over me with the idea in thought, 'Well, I'm getting a new pair of shoes.'"

## S. IV. No. 30.

(b). "In the first alternative the term 'low grade' brought back a visual picture of myself correcting papers, and not having time to prepare my teaching work effectively. This was accompanied by a very strong feeling of dislike. I seemed to anticipate the term 'future promotion' in my thoughts, and it served as a very strong reason for disliking this idea. With the second alternative came a more pleasant feeling; I had the thought, 'Well, I will just have to be careful with certain comforts,' and a vague image of a schematic table as representing the word 'food' which came in verbal motor terms, cropped up. This image came up clear for a moment and then faded, and I knew that this was what I was going to do—to economize in food. The idea of the second alternative persisted, as it were, under the surface, as if part of me. After the decision I felt very relieved and pleased with my choice which seemed a very good one."

## S. V. No. 8.

(b). "The 'dining alone' idea passed completely out of mind. Then came a visual image of dining out with a friend. We went into a restaurant and before us were the tables with white cloths and flowers in the centre. This brightened and seemed to fill the whole of my mental consciousness. There came a decided feeling of relief, the strong tenseness in my throat went away at once, and I seemed to breathe deeply again."

## S. V. No. 25.

(b). "After the first sentence I had immediate visual images of Colleges X. and Y. and I felt at once I would choose Y. The real decision came with the fading of X. and the

fixation of Y. and the reason came, 'I can attend what classes I choose at Y.' At once tension disappeared and there ensued a decided feeling tone of satisfaction, and I settled back in my chair."

S. VI. No. 1.

(b). "The image of a small racing car with a yellow body came and persisted. I did not consider the blue. I had the idea, 'Well, the yellow looks nice but the blue is characterless; I'll take the yellow.' I had a distinct feeling of satisfaction, and the image of the yellow auto grew more distinct and then gradually faded. A feeling of myself possessing the car seemed to take its place."

S. VI. No. 3.

(b). "There was a slight conflict between the thought of betting on a certainty and the pleasure of the bet itself. I thought, 'Well, if the other is foolish enough to bet I might as well take it. The bet would make for partisanship.' There was an increased feeling of tenseness and excitement felt throughout the whole body as muscular tension. The feeling tone was a decidedly pleasant one. There was a kinaesthetic feeling of being personally absorbed by the image 'of myself making the bet with my cousin,' that constituted the decision."

S. VII. No. 6.

(b). "The idea came, 'I'll ask a policeman for the carfare.' This came first and stayed there. Then came the visual image of a policeman on the corner. He was fat and swinging a club. The thought came in then, 'It's too far to walk,' and here the visual image became richer, more elaborate, and more vivid as the decision formulated. Then it gradually faded, and left a pleasant feeling—one of satisfaction. There was scarcely any tension throughout; it was merely following out the first impulse, and there was no conflict."

S. VII. No. 18.

(b). "The idea of my last illness—an operation for appendicitis—came to me, and I had a visual image of myself after being in bed for a long time, with the wound refusing to heal. Then for a moment came the pain image—the pain in the back that followed—and I remarked myself saying with my lips, 'anything before that again,' and the image went out. The strain relaxed and some slight satisfaction came then, though a slight feeling of disagreeableness still remained."



## S. VIII. No. 9.

(b). "After turning from one to another, my gaze became fixed on the sleeping one,—that on the left. The thought came that it seemed both softer and sweeter. It looked more natural, while the other was more fixed and glassy in stare, without the softness of the sleeping face. A decided feeling of possession and familiarity came at this point with the one decided on, while the other passed completely out of my mind, as if it was not there. There was an agreeable feeling tone throughout, and only a slight degree of tension located in the forehead, and the breathing seemed to be light until after the decision."

## S. III. Series II. Round 3.

"There was a feeling as of familiarity due to the olfactory image (oil of cloves) of the green liquid. This attracted me much more than the purple bottle, to which I appeared indifferent in comparison. There was a definite act of acceptance which seemed to cause me to kinaesthetically turn away from the purple one as if it did not exist, and to go towards the green. The decision was accompanied by a sudden liking for the one and not for the other, which made the one I chose stand out very clearly. Marked satisfaction accompanied the choice."

## S. V. Series II. Round 5.

"There was a feeling of satisfaction in perceiving the two that I knew I liked, and a verbal motor image came with the words, 'These are the two at the top end of the scale,' I was vaguely aware of how they smelled, yet could get no definite olfactory images. I merely fixated the one I liked best and which seemed to have a more pleasing note, and I seemed to say again with verbal motor imagery, 'I'll take that.' The other one seemed to recede from view into the background as if withdrawn spatially, while I fixated the chosen bottle with satisfaction."

## S. VIII. Series II. Round 2.

"On perceiving the red I immediately passed over to the green which I knew I liked better. I could not recall its smell, but remembered its quality in meaning, and that the red was not nearly so good. I also remembered in meaning that I had been disappointed in the red before and that its colour was deceptive. I therefore decided on the green by turning to and fixating it, and then I seemed to lose the red from the

field of attention. I accepted the green with a strong kinaesthetic image as of going over to it. There was a satisfaction in this choice, and I felt that I had chosen well."

*Conflict.* The type of decisions that falls under this head is very different to the preceding in general characteristics. In the antecedent "associative period" there appears to be a marked tendency to pass back and forth many times from one alternative to the other, and the initial trend or impulse to choose one or other is rarely present. The decision comes slowly and with effort, and the fixation of the chosen alternative is uncertain, so that the brightening of the image or the added attraction of the chosen object is not distinctly marked as in the previous forms of choice. The choice may be attended with doubt and a disagreeable feeling tone as opposed to assurance and satisfaction: there may even sometimes occur a tendency to wish afterwards that the other had been chosen. In place of the "non-existence" of the rejected alternative and the tendency to reject it, it often persists in consciousness after the choice. As a result, bodily tension is not immediately relaxed, but may be noted during the after period of introspection. The choice is described as difficult, and the degree of assurance in it is low. Such a type arises from the contemplation of two almost equivalently valued alternatives, but more often where alternatives are disagreeable than where both are pleasant. Occasionally, however, there is found a desire for both, with an inability to choose between them, but such cases are comparatively infrequent. The chief characteristics of Conflict types are the lack of certainty and assurance, the inability to fixate the alternative chosen, and in place of the strong, spontaneous images of moving over to, or eager acceptance, a distinct feeling of effort as of reaching out and taking one, or of reluctantly accepting it. At the same time, such a process is rich with feeling and content which is, as a rule, vivid and persistent. The subject feels that there must be a conflict between his choice and the other alternative, which cannot immediately be decided, but must be left for an actual 'trying-out' process or realization of ends to make the final victory of the choice complete. Such a process of conflict often arises between a strong impulse and a rational or moral motive, with the final determination of the conflict through action, which gradually automatizes it.

## S. I. No. 25.

(b). "I thought, in very clear fashion, 'If the instruction is good that's all we need.' Then I turned my attention to the right. Then came a visual image of a piece of apparatus—I cannot describe its use, but it was particularly bright and vivid in its 'brassy, shiny effect.' With that came a definite feeling such as one has when lost. It was decidedly unpleasant. I then turned my attention to the black object that stood for the instructor—that was the choice. There was a definite movement towards the side on which the decision lay. The apparatus still remained, but dimmer and as in a vague field. I felt myself give a nod at the conclusion—it may have been sensory or imaginal, I could not precisely say which,—it meant, 'I've finished that now and done with it.'"

## S. II. No. 11.

(b). "I reflected that, after all, social customs were a small and petty matter, and that the clothes of a person did not make any difference. The thought came back that I had to decide finally, and with the thought arrived the feeling and image of how comfortable one's ordinary clothes feel. The satisfaction of this was pronounced and then I decided that I would wear these things, for, even though they might be slightly out of place, still it would be better to do that, than to come in evening dress and be conspicuous. I experienced a great feeling of relief at the decision, and yet grave doubt accompanied it. Just then there returned a vivid image of the aristocratic head and face of the hostess, and this seemed to cast a decided doubt as to the real validity of the decision. There was very little satisfaction, and this slight degree came from the thought that the old clothes would be comfortable."

## S. III. No. 3.

(b). "The situation appeared as a real dilemma....the box itself loomed up in the image process in a disproportionate way. I felt myself being occupied more with the candy than with other things. Still no satisfaction came....Then an idea came, 'He's taunting me.'...The box of candy now loomed up visually again. The thought was still strong: 'Such a bet is hardly honorable.' I managed to stifle this with the thought, 'Well, I'll distribute the candy and not take it home.' Then the situation resolved itself and cleared up somewhat. Some feeling of satisfaction came that I had solved a



knotty problem, but the tension seemed to die away and not lose itself at once as before."

S. IV. No. 27.

(b). "I thought in verbal terms, 'What can I do if I do jump in? I may sink with him and both lives may be sacrificed to no end.' Then I said, 'But if I don't?' and then came the thought, 'But how would I feel after the person had been drowned?' It was too horrible to think upon and I seemed kinaesthetically to shrink from it. Probably this was an actual sensation. I could not get away from the thought, which came as a reproach, that I had done nothing to save him. Finally I made up my mind that I would call out for help and then jump in; but I got no real decision. The feeling was one of tension and strain throughout, in the head, chest and round the diaphragm. It was so marked that I felt I must get up and walk to and fro to relieve the strain. The tendency to jump in was strong but so was the feeling of regard for one's life, for I am a poor swimmer. When the decision finally came I felt tense all over, as if preparing to dive in, but with a hopeless feeling of despair. The tension did not pass nor did the feeling of unpleasantness."

S. V. No. 24.

(b). "The two ideas came and seemed perfectly counterbalanced. On the one hand came in verbal motor terms. "Honor thy Father." I didn't want to decide at all, for either side seemed fraught with pain and injustice. Finally with a real effort I reluctantly decided to accept the latter alternative. Then came an image (kinaesthetic) of myself moving slowly towards the side I had decided on, but it seemed as if the slightest effort would turn me back to the other side. There was no relaxation of strain even after the decision, for the two sides still persisted. Even though I had decided on the one hand I kept on thinking of its injustice."

S. VI. No. 1.

(b). "The visual image came of myself driving a small auto with a long yellow body, and I felt its note was full of character. In spite of the feeling of satisfaction there came a feeling as of chiding myself for taking the more conspicuous car, but made up my mind for the car, and decided to face it out with myself in spite of self reproaches. I therefore decided for the yellow auto, but the satisfaction was marred to some extent

by this feeling, leaving itself felt by a slight feeling of displeasure and a slight tension."

S. VII. No. 23.

(b). "The image of myself making an apology was fairly clear and almost obstructed the other. It held the notion that this was the easier way out. Then the words, 'a true statement' came up visually in black and white before me and I felt I could not really apologize. The conflict went on between the feelings of humiliation on the one hand, and facing a 'mob' on the other. It was so tense and the sense of strain so strong, that I wanted to end it. I wanted to take the easy way out, but this seemed too humiliating to my pride, so I determined to face the issue. The real reason that decided me against apologizing was, that this could be construed into an admission of guilt, and I could not endure that. I decided, very slowly and deliberately, 'to face the hostile audience.' The decision was accompanied by a reluctant feeling as of wishing I had taken the other course, but a dogged idea of going through with it prevailed. Tension still remained after my decision was made, but it was some satisfaction to get rid of the nerve racking experience of indecision."

S. VIII. No. 18.

(b). "I thought. 'It's a very hard question to decide,' and felt a strong feeling of strain and dissatisfaction. Then, on second thoughts, came the idea. 'I can't return them if it involves the dismissal of an employee, so I'll give the value to charity and keep them, but when I am next in the store I'll be sure to get that employee and warn her to be more careful in the future.' This seemed a compromise, but even though I had decided on it, I could not get rid of the idea that the owner of the store would be losing the value of the goods. This persisted and prevented me from feeling satisfaction with my decision, though the strain seemed to subside gradually."

S. I. Series. II. Round 3.

"As the curtain fell there was a distinct suspension of movement with a kinaesthetic contraction and set of the features, and an awareness of a tension about the face. In verbal terms came the thought, 'Both are brown.' Then I looked from one to the other, remembering that one of the browns was not so bad as the other, and again, in verbal terms, came the thought, 'Which is the good one?' I decided that the one on the left was probably the one and the final act of my choice was

a final fixation of that bottle, with an awareness of my bodily self as being in line with it. The choice was followed at once by a self-initiated reversal prompted by the thought, 'After all the other bottle may be the good one.' However, I had made my choice, so felt compelled to adhere to it, though the thought that it was the wrong bottle I had chosen, prevented me from feeling that I had chosen well."

S. V. Series. II. Round. 2.

"The thought came in verbal motor terms. 'Here is the one I want to smell.' I felt very curious to know what this brown bottle smelled like as I could not remember it, though there was a feeling of familiarity associated with it. At the same time there was a marked satisfaction in noticing the familiar 'green,' whose odor I knew I liked. I turned from one to the other for what seemed a great many times, and I slowly came to the decision that I would smell the unknown bottle and so gratify my curiosity. At the same time, the notion at the back of the decision was, 'It is taking a risk to give up what is known to be pleasant for what may be very unpleasant.' In spite of this idea which seemed to accompany the decision, I deliberately decided for the unknown brown bottle. I gave up the greenish one with some regret, and I did not feel quite so satisfied as usual with my choice."

*Indifference decisions* contrast with the preceding types in that they show a general absence of feeling tone and content, these exhibiting a marked lack of vividness when present, which makes them appear forced and unreal in tone. Again there is, in the decision period, no fixation of the attention and no spontaneity of acceptance, but a weak sense of effort in choosing unaccompanied by strain or tension. There is no close identification of the self with the alternative, the whole process appearing "extraneous." Often the sense of "outer compulsion"—as opposed to "inner conviction"—is observed by the subject: "The choice must be made; I must hurry." "It is time I came to a conclusion." "This has to be decided so let it be this one." "I must end this so I'll take that," etc. Michotte labels this type "involuntary decisions," but though they approach an involuntary type in some characteristics, there seems to be associated with them—as will later be shown in the reversal period—some slight degree of self assertion in the final act. They are the result of a form of mental inertia



or "paralysis," and do not arise from marked preferences or dislikes but from indifference and lack of similar experiences or inclinations; as a result, their degree of certainty is much lower in the scale than the preceding type.

S. I. No. 7.

(b). "There came a mild feeling of hopelessness meaning, 'I can't decide; they are equally good' . . . Then I found myself looking back and forth and saying things in foolish terms. 'This has its eyes open, this one has them closed.' There was no tendency in either direction, my attention passed to different details, eyes, posture, hands, and expression of face as alertness and sleep. There came a marked feeling of inability to find any real inclination, but in verbal terms came the words, 'I must choose.' I finally fixed with my eyes the picture on the right, thinking verbally, 'Well, this one.' Then I wondered why I had chosen and if it was because my right hand was close to the other."

S. II. No. 7.

(b). "There appeared a very faint visual image of a cut-glass dish and of some lottery tickets. The notion that accompanied it seemed to be, 'This particular church thinks it all right.' Then I considered the alternative side by thinking of people very young or stupid, to whom getting a ticket meant getting the thing itself, and this was accompanied by faint visual images of such people. Then came in meaning the thought, that it would be right if tickets were sold to statisticians and such people who understood the laws of chance, and this was accompanied by a visual image of a certain statistician. The thought came next, that there would not be enough statisticians to make this plan a success. Then my thoughts returned to the other idea with the notion that, after all, this would be permissible, because people stupid enough to buy a ticket would lose their money in any case, while children would not have the money. If this church gave it its sanction then 'very well.' The decision had to be made with considerable degree of effort, that is to say, there seemed a certain quality of indifference that had to be overcome."

S. III. No. 14.

(b). "My mind wandered and then came back to study the two pictures, and before I knew had passed off again into a day dream. Again I focussed attention on the two pictures. Finally I decided to take the position on the right, and a reason

came up in thought to support it. 'I already have a photograph in the other position so I'll try the new one.' The whole thing was merely an indifferent snap of the judgment rather than a real choice. It was entirely lacking in feeling tone throughout, and seemed devoid of real meaning and warmth to me. The tension came from the fact that I was making an effort to attend."

S. IV. No. 5.

(b). "After some little deliberation I looked out and thought, 'It has been a wet day to-day.' This decided me. There was no real convincing reason behind it, but I just decided upon it haphazardly. There was no strain, the whole affair appearing neutral in tone."

S. V. No. 29.

(b). "The arguments seemed ranged on either side as for a brief. There was strong doubt on either side and the arguments wavered to and fro. Finally I said, 'Well, it will be a continuation of studies.' The final act took the form of a kin-aesthetic image, a going over to the one side as opposed to the other. There was but a slight relief of tension and no final satisfaction or dissatisfaction, merely a neutral state. I had no confidence at all in my decision which felt cold and distant to me."

S. VI. No. 14.

(b). "There was, throughout, a certain distaste for both, associated with a sense of offended modesty. I felt that I had to make a choice, so decided on the lesser of two evils and took the left. The act of choice was that of being forced into having to do it whether I wanted to or not."

S. VII. No. 19.

(b). "Then came a blank, and I had to use effort to bring up any images or ideas. . . . had a visual image of the proverbial overtired sales girl. It came slowly, faded slowly, and was not bright. It had no perceptible feeling tone with it at all. I could not get a step further do what I would. It was greatly annoying to be thus prevented from going on. To end the strain I decided then to return the goods. The decision brought relief since it ended my amazing inability to do what I wanted to do. The real difficulty was not in deciding but in arranging the situation."

## S. VIII. No. 29.

(b). "After reading through a second time I came to the decision that I would seek some money making occupation. None came to my mind however. The decision was a deliberate movement to the alternative that I chose, accompanied by an image of moving towards it to my right hand side. There was no strain or tension at all."

## S. I. Series II. Round 3.

"Clearly in consciousness I realized the weakness of both motives and the failure of either bottle to stand out in intention. In visual terms I seemed to represent the two bottles as two marks, one as chocolate and the other as yellow in colour, the line being represented as vertical. The process appeared to be a sheer intellectual counting of these motives rather than a strong choice well motivated. The whole process was at a low level of tension as if spread out and thin in affective tone."

## S. IV. Series. II. Round 2.

"I tried to recall the smell but failed; my thoughts seemed led away and it required an effort to bring them back to the matter in hand. I was conscious of what I was doing and found it required a very strong effort. I had a voco-motor image of saying, 'You are required to do that,' with an attempt to fixate and compare the two bottles. Though my eyes were fixated,—with which came sensations of strain, yet I could do nothing; no process seemed to take place. I looked at the colour and then thought of the odors, until finally I thought, 'Well, I'll take that one; it looks as if it might smell fairly good,' yet I was uncertain of my decision and felt as if I was trying to force myself to decide on one without a real reason."

The "*Judgments*" also are lacking in any marked emotional or feeling tone, in any sense of effort, or the finality of decision which marks the Preference type. The process is smooth, even, neutral in character, and quite extraneous to the "self." Any personal preference is wont to be suppressed or excluded, the subject deliberately placing himself in the attitude of an impartial judge. Consequently all warmth and enthusiasm is markedly absent. It is as coldly intellectual as a problem in addition. The only tension experienced arises from slight attention strains located frontally, or in the case of S. V. in the throat. Once the subject's arguments have been ranged for the "pros and cons" then an unimpassioned judgment is deliv-



ered. There is no spontaneous kinaesthetic image of acceptance or rejection, but a deliberate tendency to attend to one side or the other; no feelings of possession or identification with the favoured alternative are experienced, but only the deliberate act of attending to, and fixating it, in place of the other. In respect to the final process such types have much in common with the third type of decisions, but in the whole process and vividness of content, they resemble the first, yet are wanting in that personal note that characterizes the Preference types. While they are strongly associative in character they may, as in the present examples, omit any specific reference to the self, bodily or historical, i.e. in regard to kinaesthetic reference or to intimate personal experiences. That is to say, that, in the latter case, while they have definitely been part of the subject's past, they have not been distinctly of that definite character that marks them off from the world as individual, rather than general, they only exist as symbols of thought processes instead of real and vivid purposive experiences.

S. I. No. 21.

(b). "The decision was merely a fixation of the word 'Washington.' Accompanying this was a feeling of uncertainty and dissatisfaction. There was a marked lack of personal feeling here. At first there did not seem anything missing, but on review it was different, it was lacking in a strong personal element—this 'weight' was absent. There was no specific reference to my own experiences, and no definite settling of the self into part of the situation as a recognized part of my life. The whole was quite cold and unemotional the only strain being a slight tension in the head due to inner fixation."

S. II. No. 16.

((b)). "The thought came in meaning. 'The thing to decide is, which is the more popular,—which people do I like best?' Then I asked myself the question, still in meaning, 'What are the criteria of popularity?' I decided that these would be, (a), the hearing of quotations and, (b), the popular taste in regard to their length. I decided from both criteria that 'The Village Blacksmith' was, on the whole, the more popular. In this latter part there was very little imagery and no feeling tone, and the decision came from a rational consideration, I myself did not seem to be personally involved, or to have any purpose in this. Personally it was a matter of indifference, just a case of

saying 'yes' or 'no.' I seemed to 'think' this situation rather than have a part in it."

S. III. No. 16.

(b). "The decision come from inner compulsion rather than on intellectual grounds, yet there was a reluctance to give the vote to Hiawatha; it was as if I were to be asked if 'As You Like It' were better than 'Twelfth Night.' Yet it did not evoke a personal, so much as an intellectual warmth. It was rather like casting one's vote for a candidate than making a decision, since I did not act, but merely judged it."

S. IV. No. 16.

(b). "There came a conflict between my own opinion and that of the majority, that is, between popularity and personal preference. For this reason there seemed an absence of anything like a real decision; I merely had to report what seemed to me would stand as the opinion of most people. Any feeling that came was due to the conflict in suppressing my own preference."

S. V. No. 22.

(b). "I began the work of comparison and placed Washington on the left hand and Smith on the right. I then got the headings tabulated on either side of a definite line from the images. After calling up the images I glanced down the lists on either side, and either was present at a turn of the head. The facts seemed weighted as necessary, e.g. 1, 2, 3. Washington had the greater sum total; it came to about 9, while Smith's total only added to 5. There was a definite notion of these values rather than a definite image of the distinct digits. Then the image passed out, leaving me with some slight degree of satisfaction at having made a neat solution of the problem. All the mental states were indifferent and entirely without feeling tone. There was no effort in myself but merely a cool reasoning process."

S. VI. No. 16.

(b). "Thus the image of Hiawatha persisted kinaesthetically through its rhythm, and visually through these scenes; The Village Blacksmith appeared as a living personality. They alternated in front of me, first one and then the other, till finally I decided on The Village Blacksmith, and the image persisted for a while. This was reinforced by the thought that adults would enjoy this poem more. There was a pleasant feeling

tone throughout with but a slight effort in coming to the decision point. This process seemed less personal and more of an intellectual character than the preceding. It seemed like a pleasant dream as of thinking back over old friends, but the decision come just as a mere deduction."

S. VII. No. 21.

(b): "Then I came back to the first question as to which was the more valuable, and decided on Electric Lighting. There was a slight mental strain at first since the question seemed indefinite and I had to decide the conditions. Once these were settled the rest was easy. The whole process occurred outside myself as it were; I did not appear in the picture. There were no intimate feelings and the satisfaction at having come to a decision was very weak."

S. VIII. No. 28.

"Then an image of a schoolroom came, which persisted and stayed till the decision was formulated. It did not end conclusively since I felt I had no proof. I was satisfied with the decision in that I answered as well as I could, but was not satisfied with its finality. The whole process seemed to be going on outside me without any participation in it at all. There was no pronounced effort to speak of, and only a slight tension in the head at the beginning, which passed as soon as the images began to come."

Though a few extreme cases of the types naturally tend to approach each other, the greater bulk of the introspections establish the characteristic differences as sufficiently and clearly as the selected examples. It is now possible to summarize and examine the groups as a whole, and, by a process of comparison, to indicate the essential distinguishing factors of each type. These occur, first on the associative side, and second, in relation to the immediate factors of decision. The former is not only evident through the material content, but through psychological content as well. Thus, while the first and second types yield a vivid personal material content, in which the present situation is approached through a warm and intimate set of personal experiences, the last pair are wanting in this respect, the only bond being a coldly impersonal and intellectual one. The former appear full of strong emotional disturbance, the latter only arouse the tension due to the process of fixation in attention.



It is in the final process of decision, however, that the differences partly indicated by the "associative" stage seem to come to a head. That they are not due to the pure mechanism of association is evident from the fact that, in the examples used, the fourth type is of the same character and yet is wanting in that warmth and intimacy that accompanies the first and second. On the other hand, the first and second types differ in degree of satisfaction and confidence also; the first is calm and assured, the second is tossed to and fro, and the final degree of assurance is not great. In the third type confidence is almost wanting, while in the judgment type the reaction is of a still different nature. Assurance, in the last named type, arises purely from external associative factors; if the facts are thus and thus, then the deduction follows, and the assurance comes, not from a feeling of self adjustment, but from a knowledge of certain extraneous facts that merely add themselves up under supervision. In such a case, personal opinion or preference is sacrificed to what appears to be that of the majority, but in the case of the preference types, where strong emotion is aroused, the subject is prepared to maintain his preference against all comers. Thus the judgment type is open to the influence of further arguments, pro and con, but the preference type is almost beyond the force of intellectual arguments and reasons; the former is impersonal, the latter of a real and personal character in its finality, its ultimate factor being subjective and ego-centric. Its effect on the object of choice is to enhance its position in the focus of attention—as mere associative factors of themselves cannot—either through the visual or the ideational forms, and to reinforce them with reasons appearing through the conceptual side of thought processes. Obversely, the tendency is to put the rejected alternative out of attention, or to disregard and disparage it, by derogatory "reasons." This selective effect may be summed up in the term, "purposive," which includes, as a necessary condition, the historical past of the individual, while, on the other side, there is a bare reference to some extraneous facts that have no bearing on the past of the individual, so far as it affects his present line of action. While the act of judgment appears complete in itself, on the other hand, a voluntary act of choice is a preliminary to future action, first a tentative trying out, then a decision culminating in a purpose which implies a determination to

follow out the projected course. This purpose is developed into a definite set, which is more or less fixed in character, and which, on the receipt of the proper cue, will tend to consummate itself in action.

The problem then is to explain what may be the factor that is present in the act of decision as the determining agent in purpose, and evidently lacking in the case of judgments. Michotte and Prüm assert it to be "the self" of Ach's "determining tendency;" Wheeler, on the other hand, finds that it may be ultimately reduced to definitisation and localisation which finally passes into a kinaesthetic image. The results of the present work agree with that of Wheeler in asserting that the final act of choice results in a kinaesthetic image resolving into a definite mental set, and signifying "acceptance of the chosen alternative," but cannot sustain his final conclusion that this is a complete and sufficient account of the matter. There still appears one factor unaccounted for, viz. the agency that determines this set, and in this regard this investigation would tend to follow Ach, and account for it by some prime psychological factor such as the "self."

While, on the one hand, it might appear as correct to assert that "the self" takes part in the judgment, as it is to say, that it takes part in a decision, since in both cases the "I" appears to be reported as the agent, at the same time there is a marked difference in reference; in the one case the decision is ego-centric in character, but in the other case, impersonal. The one implies the assertion and accompanying determination, "This is what I, *myself*, will really do," the other, "This is what appears to be really correct in so far as I (or *any one else*) am able to interpret the facts known to me." The difference is brought into relief by successive experiences of decision and judgment, the latter experience tending to elicit a note of surprise from the subject as to its ease and the lack of effort that marked the process. At the same time, the fact of consciousness of self content in relation to a situation does not constitute the essential difference, for it may accompany an act of judgment as when, for instance, one attempts to compare one's person, conduct, or character, with those of another, or, again, as a concomitant of an introspective report. On the one hand, some amount of self reference and self content invariably accompanies a decision, and culminates, as judgments do not appear to do,



in this specific act of self-assertion, implying determination to carry out the projected purpose. The manifestation of the self-assertive tendency has been described and developed by McDougall (11, p. 67 & Chaps. VI. & VII.), and may be regarded as identical with the factor that functions so strongly in Ach's "determining tendency," which, when provoked by lapses, produced the characteristic intensification of resolve through the self-assertive expression of "I can and I will." That this tendency may act as an accentuating factor in attention capable of effecting the change at the decision point, particularly noticeable in the preference types, is postulated by Pillsbury (16 p. 255-256), and is experimentally demonstrated in the work of Morgan (14), who found that interference stimuli increased the degree of concentration of the individual upon his task, and tended to produce better results than when the work proceeded without, though at a greater cost of effort. Gates (6 Chap. IV.), through an investigation of the result of recitation work, as opposed to work learned only for final testing, showed that the recitation of memory tasks was also productive of improved results, some of which are as likely to be derived from the social stimulus of "recitation," as from the factor of intelligently directed "focalisation" of attention. It follows, that the degree of self-assertion that might be expected to be present would be greatest in the case of the preference type, less pronounced where conflict was present, and weakest, or even absent, in the case where the decision was indifferent, or where only a judgment was demanded; the results of the succeeding section will show this to be the case.

(c). *The Reversal Period*

In the case of the after period the thesis put forward to explain the "final moment" of decision may be demonstrated by interference methods. Normally the decision, after being focalised, passes out of the central point of consciousness, but leaves its trace in some form of disposition as a preparatory mental set or readiness to act, the intensity of such a set being dependent upon the intensity of the self-assertive tendency. The proper cue for action will touch it off at once into *realization*,—it has, in this respect, developed into a "determining tendency." If, now, this tendency be interfered with it shows some of the characteristics demonstrated by



Ach. The method used in the present investigation was to ask the subject to reconsider his decision with a possible view to its reversal. Immediately there was set up a conflict between the acceptance of the new task or Aufgabe, and the definite set, the product of the self-assertive tendency, previously developed by the decision. It is evident that there has been brought into the field another factor that, by the way, has also been noted by Ach; this is the nature and degree in which the task has been "accepted" by the subject. In this bare act of acceptance there is a similarity to the process commonly known as suggestion, and both may be accounted for by the operation of McDougall's submissive tendency (pp. 67, ff., Chaps. VI. and VII.) In general, the act of submission is reinforced by the act of self-assertion expressed as, "I will do this task that I have accepted," though the one does not always necessarily imply the other. In the case of Ach's work (2), a newly reinforced self-assertion, caused by a lapse of effort, tended to intensify the "set" induced by the submissive tendency; in the present stage of this investigation, acceptance and previous determination, i.e. the operation of the submissive and the self-assertive tendencies, will be found to stand in opposition to one another.

If the degree of the submissive tendency could be considered in the case of each subject as a constant, the strength of the self-assertive tendency might be measured up against it, as to whether the struggle in its favour is brief, and decided against it, protracted and vacillating, or prompt in yielding. As a matter of fact the strength of submission could hardly be regarded as quite constant throughout. Different types of subjects, different moods of the individual at different sittings, familiarity and consequent mechanisation, tended to produce changes in its force, but, on the whole, the changes were not so great as to invert the relationship at any time, so that, in each case, it may be regarded as a rough individual ratio or coefficient, indicative of the strength of the self-assertive tendency. A detailed study of the after periods of decision types will bring these factors into prominence.

The *preference* type generally appeared to yield similar results to the process of decision. The final result was, in the main, a hearty rejection of the suggestion to reverse the decision, and reaffirmation of the original choice. As soon as

the subject apprehended the contents of the paper demanding reversal, irritation and annoyance were felt, later degenerating into a settled "hopelessness" in regard to the task. The immediate tendency was promptly to reject such a proposition, and to reaffirm the fact that the decision had been arrived at and could not be changed. However, an inhibition of these initial responses and an attempt to conform with the demands of the task followed, though still accompanied by feelings of futility and irritation. Then commenced the struggle previously mentioned, the submissive tendency tending to evoke a striving to recall the rejected alternative, and the self-assertive tendency striving to suppress it. As a result, though the time period was protracted, the subject was able to report but little. If the image returned in a favourable light it only came slowly and by degrees, to be quickly replaced by the favoured alternative, or if it persisted, the result was that a favourable turn to reasons for its acceptance might be developed into some reason derogatory to its acceptance. On the other hand, it often appeared with its disparagements even more intensified than in the main period. The result of such conflict showed its effect in stresses and strains; the mental effort involved being greater in such cases than in the original main period. Most subjects, when asked to represent graphically the result of such effort, gave a representation somewhat like Fig. 4, the original strain being gradually increased in intensity until further effort reached a plateau, when an augmentation of the self-assertive tendency generally resulted

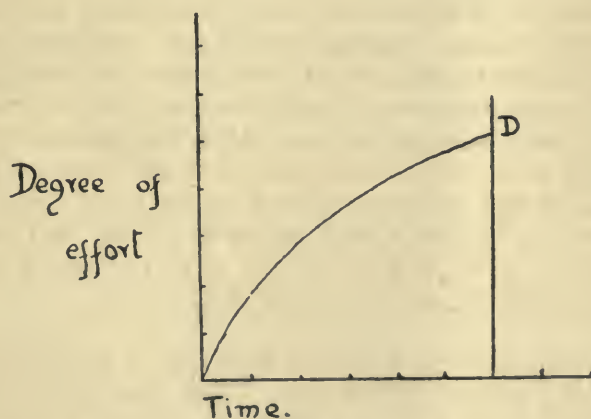


Fig. 4

in a final victory in its favour, the tension being markedly released, satisfaction being enhanced to an even greater degree than before, and assurance being strengthened by this prevailing factor; this result arose from the fact that the subject had "had his own way," i.e., had gratified the self-assertive tendency. At the same time, along with this strengthening of self-assertion, a similar phenomenon of intensification of the alternative, accepted or rejected (according as the preference was negative or positive), occurred as in the main period.

The psychological effect of such a conflict upon the subject was felt to be markedly unpleasant in each case, and the mutually inhibitive result of the temporary deadlock only to be compared to a form of mental paralysis. "Nothing happens," the subject repeatedly reports, "but an expenditure of sheer effort against an impassable barrier," the whole process of mental tension being accompanied by physical strains made evident, even to the experimenter, through external "behaviour" effects of bodily and facial contortions. This experimental phenomenon appears to be representative of the typical conflict productive of psycho-neuroses. If the acceptance of the Aufgabe be expanded in importance to the magnitude of an individual's acceptance of social convention and tabu, and the chosen alternative, with its addition of self-assertion, be magnified into the terms of a consuming desire, there is a possibility of an indefinite prolongation of the conflict, with similar qualitative results. One other symptom noticed among some of the subjects, especially in the earlier stages, was the tendency to avoid conflict by a substitution of certain factors in the original situation so as to comply with the demands of the task. In this action may be traced the method of substitution for reality, which the mental adjustments of certain psychotic types tend to adopt.

S. I. No. 20.

(c). "The first thing noted was again that attitude of hopelessness in making a reversal. There was a period of definite inertness, an absence of any activity. My first impression was, 'I can't reverse this.' This was then followed by a deliberate effort which took the form of opening the cover to re-examine the pictures. The idea at back was, 'It's impossible but I'll try it.' That was followed by repeated alternate fixations of the two pictures, one of them being quite



divested of feeling of any kind—a purely cognitive apprehension, the other having a definite feeling tone to it as being ‘my property.’ This was the same feeling that one gets in connection with the attitude to one’s brother as contrasted to one’s associates. The one is somehow identified as being yours—as part of you, the others do not count. Again I ceased to attend to the other alternative and thus continued fixating my choice with a strong feeling of possession which was pleasurable. The muscles of my face were fixed in the one case, in the other I felt them relaxing as in an incipient smile. I cannot say whether this was sensory or imaged. The decision came back with renewed and even greater confidence than before, and I felt a decided relief in tension and a satisfaction with my choice which I still continued to attend to with the kindly feeling I designated ‘possession.’”

S. II. No. 25.

(c). “Was annoyed to see the experimenter reaching for the paper that signified ‘reversal.’ I marked in inner speech, ‘Oh dear me!’ At the same time I experienced extreme annoyance and irritation. I thought, in meaning, ‘Now I must get my images back.’ I closed my eyes to shut out visual percepts in order to recall them. With a strong effort I recalled the apparatus, which came first as vague objects. It next took on the appearance of chemical apparatus, of test tubes etc. Then there came a very disagreeable feeling tone, as some of the more vague took on the appearance of tarnished brass. At the same time the idea came with this meaning, ‘Apparatus with poor teaching is useless; the essentials are lacking and only the auxiliaries of performance remain.’ There was a general feeling of emptiness in contemplating the apparatus as against the richness and fulness of instruction. There was a final impatient rejection of the ‘apparatus proposition,’ with this inability to decide for it. The decision here took on a kinaesthetic image as of a going back from the non-pleasing alternative, which amounted to a rejection.”

S. III. No. 23.

(c). “Was irritated and annoyed at being asked to reverse this but though I started out to do it I felt the hopelessness of the task. In this situation everything pointed in the one direction; I could not think of any doubts as to what course to pursue. Even the ‘hostile audience’ which was glimpsed

visually for a moment seemed to bring no weakening. I felt I could face the outside world alone. There came an idea of possible martyrdom, the word, 'Spinoza,' came in a motor speech image, and I felt as if I could defy the world. Then the thought came tinged with strong feelings of annoyance and displeasure, 'Apologize, and you lose self respect; you would be worse off than if you opposed the world.' On the other hand came the pleasurable thought that my course would be approved by most people in course of time, so that under no circumstances could I relent, I thought, 'The decision is mine—tenaciously mine,' and I had a notion of embracing it as if with my two arms. Here I felt my hands gripping the arms of the chair. This was the decision. Immediately the emotional tension subsided, the tightness about my chest relaxed and I began to breathe deeply."

S. IV. No. 6.

(c). "It seemed impossible to make the attempt and I felt annoyed at being asked to do what I had already completed. Still I made the attempt to gather a few arguments. I tried to recall all the circumstances. I had to try hard to do this. Again came a visual image of Columbus Circle, at one side was the subway entrance, on the other I looked along Broadway uptown, and slowly came the idea, 'Well, there will be many interesting things to look at as I pass along, and I am fond of looking at shops. I will now have plenty of time to enjoy myself.' This did not sound very convincing, and I quickly returned to the idea of accosting a stranger. I tried to convince myself that this would hurt my dignity. But I felt that this was ridiculous, this accident might happen to anybody and no decent person would mind lending me a nickel for car fare. The idea grew stronger and the visual image of myself asking someone, seemed to come up very bright and clear. Immediately I felt myself again accepting this idea with great satisfaction. I felt even more confident than before that I had chosen the right course. The decision brought relief from all strain and tension, and left me with a relaxed sensation and agreeable feeling. The feeling of the impossibility of reversing was in the background all through and I felt any arguments for it were rationalisations rather than convincing reasons."



## S. V. No. 18.

(c). "First I straightened up in my chair, I felt my face flush and my lips tighten. In verbal motor speech I said, 'I could not change.' I had not expected the paper and it came with a decided degree of unpleasantness since it was so unexpected. There was both disgust at, and indignation with, the task. However, with a great effort, I attempted to recall both images and they came, but very weakly. The former, the rejected alternative, was so distasteful that I refused to consider it. The other returned with a haze in front of it, but had so strong an attractive force that I could not resist it, but I felt as if pulled bodily towards it as by a magnet. I had a kinaesthetic sensation as of moving in its direction. As soon as I had redecided, I felt a glow of joy at not being able to change and a great relief from tension, more especially in the muscles of my throat, as if I had been released from a very disagreeable duty."

## S. VI. No. 10.

(c). "I said in inner speech, 'Don't want to do it this time,' and a strong feeling of displeasure at having to do this came over me. Then came a period in which everything seemed to be a blank. There were images teeming in the background, but I could not bring them into focus. This inhibition seemed to be due to the displeasure, for, when I got rid of it after a strong effort, then I got glimpses of what I was trying to see. First came an intensified image of myself, clear this time, in summer clothes and without a coat. This faded into the background and never came out again very strongly. I then had an image of riding on the subway and the thought, 'I'll get home more quickly than by bus,' and then I forced myself into thinking that there would be six blocks to walk after getting out. I turned next to the bus, which came visually at first and I saw myself shivering on top and felt the cold down my back. This was too unpleasant, and I felt I had to let the former decision have its way. Immediately I did so, relief and satisfaction came back and I felt myself breathing again. I never even regretted having to walk six blocks from the subway in the rain, for I looked on this as slight in comparison with the bus ride. The effort to force down the original decision was very strong; at first I thought I might change, and the strain of effort seemed to move up and down like this,



(Fig. 5) with the weakening of resolution, but the previous decision was too strong to alter."

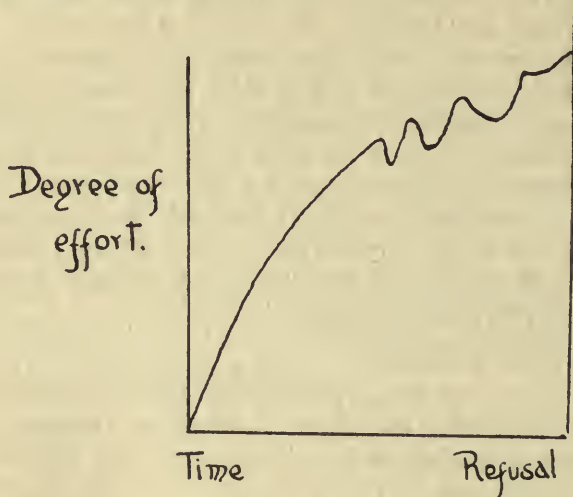


Fig. 5

S. VII. No. 9.

(c). "I thought, 'I don't like the other; I don't want to change.' I felt I had to make the effort. I went back to the one I rejected determined to find some good points in its favour. I thought, 'The child has nice eyes,' but then I noticed the word 'Copyright' across the chest. This annoyed me, and the smile or smirk was even more distasteful. The bow behind the head seemed to push the head forward and I wanted to push it back. I had a kinaesthetic feeling of tightening up my muscles as if about to do it. Then, again, I made a hasty comparison of the points of each, but I felt I disliked the rejected one more than before and my choice seemed even more preferable. I felt that no amount of effort could make me change my decision."

S. VIII. No. 6.

(c). "A 'hurt feeling' came since I felt I must go over again what I had decided on to my satisfaction. This arose every time I attempted to reverse. I started to cast aside the idea of asking someone for the carfare, and summon the image of walking home but I could not do it. I turned back to (b), and thought by trying to forget my original decision I would be able to see the problem over again as when I first saw (b). When I made the attempt, however, the original

decision with its image came up, and kept in the field of attention. I tried to think up the 'crime wave' reasons again but had to dismiss these as ridiculous. It seemed so foolish to walk home when, by the easy method of asking somebody for the carfare, I would be able to ride; it appeared both foolish as well as exceedingly distasteful. So strong did my dislike become that I gladly let it fade away, and then I tried to convince myself by further reasons, but they only appeared ridiculous. I was compelled to allow my previous decision to return, which it did very vividly. Then all the dissatisfaction passed off together with all the tension and restrained breathing. I felt greatly relieved."

S. III. Series II. Round 2.

"At first there was an initial tendency to comply with the request for reversal. But when I came to attempt it, the thought came in verbal terms. 'Nonsense, why should I reverse,' for I knew that I wanted the one on the left. Then again a thought came verbally, 'Why not? I may be able to smell the one I want in the end.' But a reply came again in thought, with the meaning, 'There is no reason to put it off; you have the privilege to choose as you will,' and a strong feeling of masterfulness came with the thought, that as I had the power I was going to use it. The overcoming of the seeming opposition felt very gratifying. Previous to the decision there was strong muscular tension from the waist up; after the decision was arrived at, the tension relaxed but I felt my jaw slightly clenched or set."

S. IV. Series II. Round 2.

"Here again I could not reverse, though I made a strong and persistent effort to do so. At least I could so far as sensory elements go, I knew that the colourless bottle was not so bad, and if there were any incentive to do so, I could readily accept it, but I could find no reason for accepting it. There was no specially strong motive in the object itself to suggest reversal, so I decided to abide by my first choice of which I felt more assured even than before."

S. V. Series II. Round 4.

"I felt annoyed throughout this attempt. I knew I could not reverse even though I tried to force myself to, for I wanted very much to see how it would smell. I had a verbal motor image of saying, 'Perhaps the "brown" that I have chosen is

a disagreeable odor and the green one pleasant,' but I did not feel convinced at all, and it made no difference to my former decision. I then decided not to reverse but to stick to my previous choice. This final decision gave me very great relief from strain and pronounced satisfaction."

*Reversals of the Conflict* type are even more complicated and unpleasant than the preceding. In the previous cases, the subject frankly acknowledged to himself the almost sheer hopelessness of the task but went about it with some showing of effort. In the present examples the subject was not always ready to entertain the prospect of reconsideration, for it appears to involve the recurrence of the whole conflict, with the rejected alternative returning as a disturbing factor, and upsetting a state which might best be described as temporarily static. Thus there was an obvious shrinking from a return of the difficulty of recalling, and then repressing, the impulse that had once been repressed only with effort. As a result, there was a marked tendency to keep it from coming into the field of attention at all and thus receiving consideration, for its potency seemed to be feared by the subject. Accordingly, the mental procedure of a reversal here may appear as fumbling and pointless.

There is, on the one hand, an implicit avoidance of this potent, yet repressed, factor, and, on the other, the knowledge that, if the previously chosen alternative is considered, then only factors that are favourable to its acceptance can be brought into the field of attention. A third alternative, that of distorting the original situation is sometimes adopted here, but rejected after recognition. As a result of the dilemma, the final tendency is to take refuge in reaffirming the original decision and rejecting the task. In this way, the confidence in the reaffirmed decision is enhanced, and the degree of assurance in it rises almost to the level of the original preference choice. With its reaffirmation there is a tendency for the good points of the chosen alternative to be reinforced, an addition in its clearness in attention occurs, and there is also the increase in confidence already noted; or, again, the choice may be negative with the rejection factors reinforced. Such a reconsideration has, in effect, two possibilities. The first is similar to making over the decision a second time with an increased, and stronger victory than before, for the chosen



alternative; an increased measure of self-assertion appears to account for the phenomenon, which is confirmed by an appeal to the subject's ratings. In the second case, where the self-assertive tendency was not strong, but the acceptance of the task was a pronounced feature, the latter becomes the chief factor in the decision, and ensures the acceptance of the previously rejected alternative, but with no great degree of confidence or certainty.

S. I. No. 11.

(c). "A current of resistance at being asked 'to do this kind of thing and vacillate and change my mind,' was at first encountered, but this was inhibited by an effort. There returned fairly readily, a revival of the two representations. The first was a visual and kinaesthetic representation of evening dress, and, while this remained, I found myself at the same time considering my person as it is at present. Then followed a process of reflection in which I thought,—I cannot say in what terms—I could, without discomfort follow the other plan, since I found myself inclined equally in both directions. But when I turned my attention again to the image representing evening dress, the judgment that, under the circumstances described, this was my decision, again returned to me. With that I seemed to pay no attention to my person, which stood for the other alternative, and my affirmation of the former decision was complete. The process was much more difficult than the preceding part, but the images were less vivid in spite of my efforts."

S. II. No. 19.

(c). "Was decidedly displeased at having to go back to a decision so difficult to arrive at. I began with a visual image of where I had left off. It came back fairly readily, in a partly visual and partly kinaesthetic form. I both saw, and experienced, the movement of taking up the goods, putting them in wrappings, taking them off the bed and walking out the door with them. The next part would not come back at first, but at last, very dimly, I saw the store and the girls employed there. At any rate, I thought to myself, I will not have to witness the dismissal of the employee, I will merely put the goods on the desk and say, 'I was given too much goods. My address is on the package. Kindly see to the matter. Good-bye!' so that I would have very little to do with such an unpleasant experience. Then the idea came, 'It is an unwar-

ranted waste of my time.' then, 'That's silly, I'll telephone,' came with inner speech. But I thought immediately, 'I hate the telephone.' I felt even less decided than before. The whole was accompanied by severe strain and was decidedly more unpleasant to arrive at, than even the original decision."

S. III. No. 11.

(c). "It took a long time for the reversal. The greater part of the time was consumed in attempting to reconsider the problem, the inhibitive effect was marked, and my mind persisted in wandering away. I had to exercise great effort—felt kinaesthetically in the back and chest—to bring myself to the point of reconsideration; I had not been quite satisfied with my previous decision, but finally succeeded in recalling and eventually fixating it. I called to mind the rejected alternative as it appeared before, but the phrase seemed to come to my lips in inner speech, 'Oh no, I wouldn't do that,' with strong feelings of displeasure at it. Then I thought, 'Well, I'll turn down the invitation,' but I realised that this was not allowed by the terms of the Aufgabe. I tried again to reconsider the rejected alternative but it was accompanied with such disagreeable qualities that I quickly relinquished it, with the thought, 'I couldn't go and chance being the only one in evening dress.' I therefore went back to my former decision and at once felt a marked relief from tension, but there still stayed behind a slight degree of dissatisfaction with the alternative. The struggle was accompanied with suspension from breathing and a feeling of flushing in the face as well."

S. IV. No. 3.

(c). "The first thing that came was, 'No more reversals.' I had to repress this idea and put forth a special effort. My lips moved and I repeated to myself, 'Endeavour,' with the meaning, 'Get over your mental laziness and make a special effort.' I felt my brow wrinkled and I put my hand to my head. As soon as I made the effort, such a rush of ideas began to come in, that my mind was overwhelmed and the whole faded quite away. I had to make an effort to recall the whole occurrence, which did not come back for some time. I had to again make an effort to recall what it was about. Then the whole picture returned with a rush of feeling. The 'fellow' of the previous part who had been waving his hand, came back as a visual image and was standing beside me. Again he appeared to wave his hand. Then followed a blank. Then I brought back



the idea of betting and finally made a rational decision. 'I could not take a bet with such odds.' Then I set the whole thing aside and refused to consider it. I closed the paper, but I had to repress a feeling of antagonism and the impulse to bet. Right through I had to keep down this impulse to accept the challenge when he seemed to be taunting me, and I was afraid I would yield to him if I let it come into clear consciousness. The whole experience was full of strain and tension, and very unpleasant, and even after making an effort to end the decision, I felt that if I gave it a chance it might return if I considered that side of it."

S. V. No. 24.

(c). "The paper signifying, 'Reverse your decision' made me feel dismayed. I had not wanted to make a final decision before and had decided reluctantly. Now the thought, 'I have decided,' came over me with a strong tendency to affirm it by hitting the table. If I let it, the whole situation would have returned with a rush, and I should have had all the difficulty of deciding over again. I felt I could not bear to face the task, so simply chose to adhere to my former decision. There was no relaxation, and very little satisfaction. I felt more certain than ever that I could not let go my decision, but all the old conflict was lurking in the background ready to occupy the field of consciousness at the slightest provocation. There was no relaxation even after this, only a degree more of confidence."

S. VI. No. 26.

(c). "I did not feel annoyed, but did not care for the idea of change. As soon as I began to reconsider the situation, the greater importance of the thing seemed to loom up and influence my choice. I tried to make up my mind to a reversal, thinking the matter only of slight importance. I felt myself saying, 'slight,' trying to assure myself that the whole affair was of no consequence, and that I could easily reverse my decision, then just at this point the idea that the matter was of great importance, loomed up. If I could keep down this idea, then the idea of a public defence before an audience would have been of no consequence. Again it came up. The conflict was as strong as if I were deciding between two vital matters. I made a strong effort to exclude it, and I felt it as a sensory motor reaction of strong muscular effort and felt myself tighten my muscles all over my body. When, finally I



succeeded in banishing the importance of the situation the former image came up, but I could not accept it now, and felt that I could not fight to accept it; the other effort had taken my strength, so that I could not reverse. There was a distinctly unpleasant feeling tone throughout, which was strong in the exclusion of the importance of the situation. I feel that, even now, all I have to do, is to let it surge up and carry my decision away with it."

S. VII. No. 2.

(c). "The thought of further consideration and reversal of the decision so painfully come at, was displeasing and irritating. The situation returned only after a very strong effort. When it returned only the image favoured by the decision came back and I failed to bring back the other. I said to myself, with inner speech, 'I will get wet; I will get wet,' for I wanted to make this alternative image as disagreeable to myself as possible to force myself to 'go over the ice.' Sometimes the 'getting wet by going through the slush' did return, to be immediately crowded out by myself crossing the ice and slipping. It seemed to be impossible to change the decision without changing the situation, and before I knew I had reformulated it. There seemed to be very little ice, only enough to be dangerous, but then the slush lessened too, and the ice then seemed to be worse than before. Before I could control it, the whole situation became even more nerve racking and dangerous, and the little storm became a heavy one. Then I realised with a start, that I had been distorting the situation and had to recall the original over again. It came with a very fleeting and indistinct image; this persisted in a hazy fashion, and my old decision, that of passing through the slush, returned, but not so clearly as on the former occasion. The notion of risk seemed to be in the background right through, and prevented me from feeling the slightest degree of assurance that would let me accept that alternative, while the getting my feet wet did not seem so dangerous when compared with it. After the decision came, some tension still remained, and though I was glad that I had not changed I did not feel comfortable in regard to it."

S. VII. Series II. Round 3.

"I knew that the odor of the violet bottle was much more pleasant than the brown, but I had been more attracted to the latter in the end though the idea of 'risk' seemed to accompany

it. I felt annoyance at an inhibition of my choice that I had regarded as logically made; still it took an effort to come back to the one I really wanted, for it seemed as if I were forcibly putting by the attractive one, and taking the one I originally chose."

S. II. Series II. Round 3.

"I felt it possible to reverse by giving up my previous choice, that I was still uncertain about. I had to overcome my mental set by sheer effort. It was not pleasant and took a long time. I had to force my attention to remain on the other alternative, and put the 'fruit juice' bottle out of my mind. I finally came to a reluctant decision to 'reverse.'"

There is little to be noted in regard to the attempted *reversal with "Indifference."* The process is one of indifference because, as in the original decision, the situation is only accepted in an impersonal fashion. Almost exactly the same result is to be found in the succeeding attempts at reversal. The only reason for non-reversal here, is conditioned by the fact, that some slight degree of self-assertion is aroused, and is somewhat stronger than the subject's acceptance of the task, but this measure of increased confidence is not consistently perceptible throughout in its results. The decision once having been made, the subject declines to change, largely from the effect of what he describes as mental inertia, or, perhaps, really from a slightly aroused self-assertive tendency. If this is present at all it does not reach the proportions of the feeling aroused in connection with the two preceding types; the reagent never mentions feelings of "possession," "acceptance," or "attraction" in connection with the chosen alternative here, but merely indicates that this particular one is chosen, not from any motives but from the necessity of having to choose. On the other hand, a common tendency was to allow the influence of the new task to prevail and thus effect a reversal, but the change was not achieved without some slight degree of opposition. The "consciousness of the Aufgabe" appeared to function as an implicit notion that, without any apparent basis for maintaining the original alternative chosen, the reversal ought, under such circumstances, to be made. Where reversals were effected this idea of "oughtness" prevailed over a weak self-assertive tendency.

## S. I. No. 3.

(c). "This part was very vague and weak throughout. It seemed to be merely a review of the whole scheme of directions to left and right without any feeling of pull or alternation. The only other thing that appeared was a feeling of ease or freedom and a sense of relaxation, and, in some way, in verbal terms, was spread the thought, 'Oh all right.' The discussion was a verbal one with no clear representation of the alternatives beyond the bare schema of direction. In the background was a whole mine of meanings. I seemed to review many points for consideration with nothing tangible to represent. I had a notion that I had simply played with the attempt and that time was being consumed, so that I merely acquiesced in the reversal. No feeling tone or bodily tension occurred at all."

## S. II. No. 7.

(c). "On observing the paper that meant 'reversal,' it appeared quite a simple matter. All I had to do was to recall the circumstances and say to myself, with inner speech, 'It doesn't really matter, one might as well hold the lottery as not.' Then came images of a cut glass dish and of people dispersing, which meant that they were going away since they had been told that something was not going to happen. The decision to reverse came very easily and with a minimum amount of effort. The feeling tone was neutral throughout and the matter gave me no concern in any way."

## S. IV. No. 9.

(c). "When the paper was handed out no feelings of annoyance seemed to come at all, there was almost indifference. Then with inner speech I said to myself, 'Suppose I change.' and the thought came, 'If the other was chosen then it is quite as good as this,' but then flashed into my mind in thought an answer, 'Keep the one you have chosen and save the trouble of changing.' I pondered this indistinctly for a while but decided that I could not change as there was no reason to do so. There was no feeling of strain right through, and very little effort. I could readily have reversed my choice as far as the pictures went, but I wanted to be consistent and keep to the choice I had made."

## S. VI. No. 14.

(c). "There was no inclination for either and no strong dislike to considering the reversal but rather a 'bored feeling.'



When I came to reconsider the matter I did not seem able to think up any reasons but just a strong dislike to change. I had tried to make a choice and now could do no more, so that I felt as if I did not want to consider a change in decision; it would cost a slight effort, for which there were no attractions to make it work while, and I did not want to overcome the inertia."

S. VII. No. 20.

(c). "I felt myself sigh as if about to make an effort, but no surprise or annoyance came. There was only a feeling of indifference. I turned to part (a), to make sure that nothing had been omitted. I found that I had overlooked nothing, so turned from one to other of the pictures, and attention passed to the one I had chosen before. It seemed to attract me a little as the face was, perhaps, clearer than the other, which seemed vague, but, on the whole, I remained almost indifferent. I did not try to make a real attempt at reversal but just permitted my former decision to remain. There was no strain, except a slight one of attention in the front of the head, but no feeling came with it; rather a passivity throughout the whole situation."

S. VIII. No. 11.

(c). "I smiled at (c), and felt no annoyance for I thought that perhaps I might be able to reverse this, since my decision was not a very confident one. I returned to the previous page, (b), but nothing came except its meaning, and nothing definite regarding it. My mind just appeared to be a blank, and I thought that I might as well allow my decision to remain for I could find no reasons to change. There was only a slight effort to try to call up something that would help me decide, but there was no response, otherwise the whole decision was just an impersonal thing that did not affect me at all."

S. I. Series II. Round 2.

"Nothing occurred except a relaxing of the face into a satisfied attitude, and a visual fixation of the other bottle. There was no apparent reluctance, and the only other thing present was the idea that neither was known to me.

S. VI. Series II. Round 2.

"I knew I could reverse by focussing my visual attention on the one I wanted to choose, so I attempted to turn to the thick brown bottle. My new determination to accept it was

not strong enough at first, and so I returned to my original choice. I could see its label through the light brown liquid, and this attracted me somewhat, but then I realised in thought, that this was a poor motive. I had always the notion in the background of consciousness, that I must reverse wherever I could to comply with the demand. This idea now came into my mind, with the meaning that, I ought to do it now, so I determined to reverse my decision. I therefore deliberately returned to the dark brown bottle and said to myself, 'I'll accept that.' Personally there was very little reaction; I was indifferent and neutral throughout. I felt definitely out of the decision, and the whole experience was rather like a 'judgment' than a definite decision."

*In Judgments*, as in the case of the Indifference types, the subject finds no serious objections to reversal. The whole matter can be considered readily enough, with a moderately easy recall of the previous images and their accompanying significance of meaning. The procedure is that of a reopened judicial inquiry, where the original verdict can be readily modified in the light of further evidence. Should the subject succeed in adding to the rejected side to a sufficient degree a reversal is possible, but if no further evidence is forthcoming, then the judgment must remain as before. The total absence of tension, emotional tone, and attendant feeling, together with a lack of purposive reference, stamp the reversal as similar in these characters to the original judgment. A judgment, however, differs from the indifferent decision, in that the process may be one of greater interest and more pleasantly toned for the subject; another difference is the readiness with which the subject can, in some instances, bring in additional evidence without any arousal of opposition. The reason for the evenness of the process lies in the absence of the self-assertive tendency, which, as in the original judgment, is also lacking here.

S. I. No. 28.

(c). "I recognised the significance of the Aufgabe and attempted to comply by recalling the unfavoured alternative. A visual image of a church steeple came at once. I was aware that this stood for my former decision and that I must now consider the other possibility. There was no kinaesthetic element involved, such as tension or a settling down to the



task as is usual in attempting reversals. Instead, the image of the school with its grounds came back, and, as I attended to it, it seemed to extend its boundaries. I was conscious, in verbal terms, of the word 'Sunday School.' This apparent widening of the visual image of the school area meant, for me, that school might be taken in a larger sense than that in which I had formerly used it, so that the term 'school' merely meant an illustration of the broader sense one takes education to mean. I was able to reverse my decision quite easily, with the consciousness that I was using an extended connotation of the term."

S. II. No. 16.

(c). "I thought, when I saw the directions for a reversal coming, 'It will not be hard to reverse.' There was no feeling of annoyance, though, for preference, I would rather have proceeded with something new, than have gone over this ground again. It did not seem hard to arrive at a new decision, for I thought other criteria might easily apply here. Indians are always popular with adolescents; these read both 'Hiawatha' and 'The Village Blacksmith.' It is quite possible, or even probable that, since Hiawatha embodies the life of the Indians, it would make the better showing of the two poems. Again, Hiawatha lends itself to display and tableau. The image of the Indian headdress symbolising Hiawatha came with these ideas and remained; there was no effort in its recall. The whole process moved quite smoothly and without tension, and was rational, rather than emotional, in character."

S. III. No. 16.

(c). "The thought came when I saw the paper for 'reversal' handed to me, that I might be able to do it here. Both alternatives seemed to be equally disliked by me. As soon as 'The Village Blacksmith' was represented, however, a rebuttal came; I felt that the former reasons for the 'Hiawatha' choice were valid and unshaken. Then I tried again to bring up the second reason but as soon as it was formulated a rebuttal came. The result was as if the scales always went down on the Hiawatha side and nothing that could be put on the other side could bring them down in turn though I tried very hard. There was not very much strain except in the head, but the whole passed off with a peculiar relaxed feeling that was different to other cases of reconsidering my decision. I felt that, in this case, I was just reaffirming my previous decision



in exactly the same way since there was no evidence for changing it."

S. IV. No. 28.

(c). "The sight of the paper asking me to consider a reversal made me smile. I thought of it as an attempt to shift the decision. 'I might be wrong perhaps,' came the thought, but this did not affect me at all. Even if the majority were against me I would still keep my opinion, for I am very positive on this point. Then came the thought that, perhaps in terms of society as a whole religion is the more important, but I could not agree with this, and felt I had to consider that school was the greater influence. There was no strain and the whole process passed off quietly and without any feeling."

S. V. No. 22.

(c). "Felt no dissatisfaction at having to reconsider. I tried to make the reversal real, and had an image of a lecturer expounding the merits of Smith, as opposed to Washington. I saw myself listening and tabulating the reasons but could get no specific points of any weight. I thought I would, therefore, abide by my former decision. Behind the image of the lecturer was the notion that, if I could hear real arguments in favour of the one side, then I could change my former decision. There was no strong feeling tone for either side. There was no compelling feeling that it *had* to be Washington rather than Smith, and no effort but a cool reasoning process."

S. VI. No. 16.

(c). "There was some feeling of repulsion at the thought of reversal, but the choice was not momentous. If I could find sufficient reasons on the one hand then I could reverse. First the visual image of The Village Blacksmith returned, and the thought that, after all, much repetition might make it insipid. Then Hiawatha came back visually, with the idea that the rhythm and music must live for all time and the images would be forever new and fresh, and, again, that things learned in childhood do persist. If a general consensus of opinion were taken Hiawatha would get the most votes. Aside from the figure of The Village Blacksmith there is little, but, in regard to Hiawatha, its myriad of pictures and the music of its verse make it very appealing; this would tend to make it lasting. The process was one of weighing,—detracting from the one and adding to the other—and then the decision was made. Notice I do not say, 'I made the decision' but that the

decision was made. My work was merely to put in the balancing factors and the whole thing was formulated. No strain or feeling process attended the situation throughout, it was rather calm and pleasant like a daydream."

S. VII. No. 22.

(c). "I realised, in thought, on reading the Aufgabe, that John Smith perhaps had not had a fair chance, and that it would be as well to reconsider the matter. . . . . There was no strain present, but I felt myself as in a history class and reciting it, but there was no personal reference beyond this. The whole thing seemed projected in front of me, and I had no part in it but merely supplied the reasons. Except for this I seemed quite passive and simply had to record the decision."

The reversal period may be said to be one of reaffirmation; the chief point to be noted in it is the effect of aroused self-assertion. In the original choice this factor is such a constant, that it passes unnoticed under such terms as kinaesthetic images of acceptance, assurance, rejection etc., or, in the case of Michotte, as that of the "self." The fact that there enters in, at the decision point, something other than the associative factors is shown by undeniable evidence. For, if no other than these are present, then it is hard to see why the attention process should change, why the attitude of the subject should be so different directly after decision in relation to the accepted and rejected alternatives, and, further, if no other factor like "self-assertion" intervened, and the choice is, merely, the product of the task, why there should be such difficulty in even entertaining the idea of a reconsideration of the situation when the task requires it. Where the degree of self-assertion is great, the opposition to a reversal of decision is found to be insurmountable. This has already been qualitatively demonstrated, as has also the fact, that, where it is weak, or absent, a reversal is often achieved. The presence of satisfaction with the choice seems to be one positive index; there is a self-complacency about a well grounded choice that stamps it as winning the self-approval of the subject. This satisfaction is enhanced by an adherence to the original decision after an attempt at reversal. Yet since its influence is not exercised through the addition of content, but rather in the regulation of the intensity of the process, the main effect of self-assertion is to be sought for on the latter side.



Confirmatory evidence is present, in that the recalled images in part (c) do not equal in intensity their originals in part (b), and this may be accounted for by the temporary lessening of the degree of self-assertion by the opposition of the submissive tendency. As a result, there is a temporary dimming of the image, the original brightness of which is subsequently recovered in the process of reaffirmation. Why, then, is self-assertion not obviously present, and reported as present, in the original decision, it may be asked? The reply is obvious. It is always present as an implicit factor in choice just as the acceptance of the task is, and, as the choice under experimental conditions, it is assumed, does not differ from the choice under practical conditions of life, so this element is constantly present throughout, and may pass unnoticed on account of its general universality. Given occasions where it is weak, or absent as in the last two types considered, then the difference is obvious. The process of attempted reversal brings it out in accentuated form and again alters the conditions of its presence, thereby rendering it explicit. Thus, we may, by an application of Mill's "Canons of Concomitance and Variation," bring out its presence by the two means, first of the different types and second by confirming this by a process of variation of its presence in degree or quantity.

#### *d. Motivation*

While presenting situations of various degrees of difficulty, the members of the first series are so assorted in basic motives, as to offer little or no comparative basis for their differentiation in that respect. Their chief value lies in their presentation of different material of an imaginative and reflective character, thus permitting a check-up with the series containing more concrete material, that is finally consummated in action. While an exact tabulation of the first series would be cumbersome on account of their variety, the outstanding motives are worthy of note. One prepotent criterion with all subjects was that of ethical principle. In the case of one individual, this was the final moment that influenced choice, principles of truth and justice outweighing more lenient considerations. Next came the various forms of a "self-assertive" character, such as dignity, pride in one's work, self respect; even personal vanity occurred in some few cases, while desire for social approbation also figured. This type of motive



marked off, as especially "forceful," certain of the situations from others that involved only matters of individual well-being and pleasure, and seemed to add much weight in consideration of the former. One very strong characteristic, in the case of two subjects, that squared well with the experimenter's estimate of their external behaviour, was solicitude for others, both in the circle of the family and society. On a lower plane, of a non-altruistic and a social nature, motives of direct potency were the desire to possess, or "acquisitiveness," and the desire for bodily well-being, comfort, health and safety.

In the second series, the situations are more or less of a uniform nature, of approximately equal difficulty and to be judged by similar criteria; these factors offer possibilities that the heterogeneous character of the former do not. The motives, too, are limited and of a homogeneous type, hence they are comparable. Again, since the forms of choice are all unknown to the subjects before presentation, they provide an illustration of progressive experience and a development of conception of values that must correspond, in some degree, with the acquisition of practical experience and its derived conception of values, that takes place in the world beyond the laboratory. The differences that are manifested therefore do not depend particularly upon past facts of association, but upon qualities such as personal traits and characteristics.

On the presentation of the first round of the series, the subject had no accurate past experience with which to guide his choice. Since the bottles containing the odors were only differentiated in point of colouring, the first tentative principle of choice was that of colour. Degrees of positive colour preferences, as well as past sentimental associations of the colour with other objects, formed the basis of motivation; in some other cases the colour was associated with gustatory images, and a tentative scheme of values borrowed from these preferences.

(N. B. The numbers given in the examples refer to the odors. See Part. II.)

S. I.

"It (No. 7.) seemed like medicine. Only the notion was present and the meaning seemed tied up with the colour of the fluid. The word 'water' came in voco-motor terms, with an attitude of uncertainty, meaning, 'It (No. 3.) looks like water

but is probably not.' With that came a feeling of indifference and a visual image of myself swimming in the water."

S. II.

"The green liquid (No. 7) brought a thought of poison. On the other hand, when I turned my eyes to the other (No. 4), I thought of fruit juice, on account of the colour. This was accompanied by a tickling of the salivary glands and a watering of the mouth."

S. III.

"The colour of the right hand bottle (No. 7) was decidedly disagreeable. It brought back a visual image of medicine; the other (No. 1), though dark, looked decidedly more agreeable. The feeling tone of its colour was much more pleasant than the first."

S. IV.

"The other (No. 2) was so repugnant that I immediately rejected it. Its colour reminded me of blood; there came a slight visual image that seemed so disgusting that I immediately dismissed it. When I came to look at the other (No. 4), I seemed to look at it, not as a liquid, but as a solid. It appeared to change its nature to a piece of dark furniture or beautifully polished stone."

S. V.

"It (No. 8) looked darker, and I thought I'd like to try it; its colour appeared more agreeable. Besides, there came in verbal motor speech, 'cough mixture,' and I had a visual image of a bottle of that liquid that was decidedly pleasant to taste."

S. VI.

"A degree of pleasure came with a notion of familiarity with the green shade of the liquid (No. 5), and I had a visual image of the apothecary in *Romeo and Juliet*, and the ghastly green light of his room. On the other hand, the purple bottle (No. 4) brought a visual image of purple ink, with a strong feeling of displeasure for the implied smell. Then the green brought a visual image of peppermint sticks."

S. VII.

"They (Nos. 4 and 7) called up images of laboratory stains at once and I had visual images of eight saucers together with a kinaesthetic image of putting specimens through these. But I thought, 'These are not stains.' Then my mind shifted to purple ink as I looked at the right hand bottle (No. 4).

The other, the brown, then looked like cough mixture, and this brought a feeling of displeasure, with a slight image of its taste."

#### S. VIII.

"As I looked at them their values seemed to come in terms of taste images. The purple (No. 4) looked like dye or ink and the thought of poison came into my mind. Then I looked at the other (No. 8), and it seemed harmless like a tonic or stain for slides. When I thought of 'tonic' an image of a bottle that I had to take as a child, came up visually for a moment.

"I could not get rid of the idea that the stuff (No. 6) was there to taste, and this brought a strong feeling of annoyance."

After this primary period of floundering had passed, and some of the values had been experienced, curiosity became a very potent factor, and persisted till the process of learning was complete. At first, the curiosity was somewhat crude, rather of the nature of inquisitiveness, and arose from a desire to know what the contents of the unexplored bottles were like. Later, it also persisted but from a higher motive—the desire to "learn and fix" the precise value. The subject frequently had a notion that a certain bottle was definitely unpleasant, but desired to identify it again, and give it a precise location in the olfactory scale. This motive was strong enough to predominate over the actual values themselves and, in three cases, by withholding No. 1 (asafoetida) after it had been chosen in preference to another of slightly higher value, it was deliberately chosen in preference to Nos. 4 or 5 (Vanilla and Oil of Cloves). The motive of curiosity was potent until its utility had been eliminated by experience.

#### S. I.

"This was followed by the realization that I had seen both but smelled neither. Then came a definite feeling of piqued curiosity with the knowledge that I had chosen the right hand one before (No. 6), but had not been allowed to smell it, while I remembered smelling the green one (No. 4). I was drawn irresistibly to the right hand and, in verbal terms, came the thought, 'I *will* see what it smells like.'"

#### S. II.

"Both appeared very dark but the right hand one (No. 6) had a sediment round the top. Then came the thought, 'coffee,'



and it interested me very much. I now thought, 'Here is a theory that I wish to test.' As my curiosity grew I felt a strong desire for it and decided on smelling it."

S. III.

"Though I was fascinated by the green (No. 5), the brown (No. 7) looked like an old friend. A warm glow came over me. I felt a strong curiosity and the thought came, 'I must find out what it smells like.'"

S. IV.

"I did not care for the colour of either and, so far as I could remember, I did not like the odor of the right hand one (No. 8). Suddenly the idea came, 'Why not choose the one I don't know about? (No. 6.)'"

S. V.

"I thought I would like to try it (No. 3), in order to satisfy my curiosity."

"I recognised the dark brown (No. 8), and had a definite idea that it was distinctly unpleasant. At the same time, though the other (No. 7) 'looked safe' I thought, I had no definite curiosity to examine it. I had an idea that it was much less unpleasant than the dark brown (No. 8), but chose the latter to learn definitely its value.

"A positive feeling of annoyance at not having been allowed to find out before came over me, (No. 1 had been withheld by E.) and, in verbal motor terms came, 'I'll find out this time if not before, I *will* smell it.' The value of the odor had no significance; I had to satisfy the feeling of curiosity."

In the after period S. reported: "Despite the fact that it had a very disagreeable smell I was glad I chose it and I thought, 'If this were not an experiment the odor would not be so bad.'"

S. VI.

"Had a strong desire to see what was in that bottle (No. 2)."

S. VII.

"I knew that the green bottle (No. 5) was pleasant, but the desire to smell the purple and find out what it was like was so strong that I could not resist, and I had a kinaesthetic image of being attracted bodily to my choice."

S. VIII.

"The red (No. 2) looked very attractive, much more so

than the green (No. 5), but I had been keenly disappointed in judging by colour before, and I strongly wished to smell the green. I accepted the latter with a kinaesthetic sensation of moving bodily towards it, while the red was lost to my attention."

The concept of olfactory values was a matter of gradual development, and formed the final basis of choice in all but one case. This is shown by establishing the correlations obtained by the Pearson formula, in ranking from previous experience after the second and third sittings, with a final adjudication by getting the subject to range the bottles in order of merit for odor alone.

TABLE I.  
EXPERIENCE VALUES CORRELATED WITH FINAL ORDER OF MERIT.

Subject.	2nd Sitting.	3rd Sitting
I	.78	.98
II	.98	1.00
III	.71	.98
IV	.84	.33
V	.88	1.00
VI	.98	1.00
VII	.98	1.00
VIII	.98	1.00
Av.	.89	.91

There is a noticeable sex difference here that is inconclusive, only because of the small numbers of the groups; this is the fact, that the judgments of the women subjects were closer to final values even at the second sitting. So far as these numbers show, the women had a keener sense of olfactory values than the men, and arrived at these much earlier. Where consistency is concerned, the women's judgments agree much better with that of the group than do the men's. S. IV. exhibits a peculiar perversity, in that his final judgments are less in conformity with his sense of olfactory values than his second choice. The explanation lies in a certain degree of prejudice on account of colour and dislike to what had been previously rejected entering in the warping the judgment.

In the earlier attempts to identify the odors the subject endeavoured to recall, by means of an olfactory image, the odor that belonged to the bottle concerned. At first, there was marked inability to do this, but, at the same time, the subject was able to make a judgment by means of what may best be described as bare awareness; here there is no clear consciousness as to reasons, but only a dim perception followed by

strong conviction that one odor is preferable to the other. Later, olfactory images were used to a slight extent, and finally, a recall of relative values rather than a recall of the actual odors themselves. The desire to "play safe" and choose a pleasant odor, where such was recognised, as against an unknown bottle of sinister appearance was, at times, stronger than curiosity itself. The values that were most slowly acquired were those in the middle of the scale, while of those at the extremes, the more disagreeable were not learned so quickly as the more agreeable. Even in the final period there was not always a perfect correlation of choice with order of merit for odor, and these mistakes fall, with the one exception of S. IV, about the middle of the scale. Fig. 6 shows the distribution of these errors according to the misplacing by the subject; *a* shows them at the end of the fourth, and *b* at the end of the sixth round, of the series.

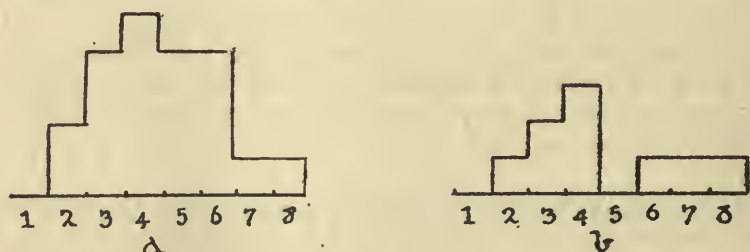


Fig. 6.

Graph showing Collective Distribution of Errors, according to Position in Scale.

### S. I.

"Then I looked from one to the other remembering that one of the browns was not so bad, and in motor-verbal terms came, 'Which is the good one?' I had almost decided on the darker (No. 8) as the good one, and had visually fixated it in order to reassure myself. This was followed at once by a self-initiated reversal. This was prompted by the thought, 'After all the other bottle (No. 7) looks like the good one.'"

### S. II.

"There was temporary fluctuation from one bottle to the other, with the thought in the background, 'Which one do I know?' Then attention and interest were focalised on the white one (No. 3). The thought came, 'It looks "innocent," but last time an innocent looking bottle was very disagree-



able.' A feeling tone transferred from the previous experience to this one, caused me to turn my attention from this with a kinaesthetic image of repulsion from it, and I passed to the bottle on the right (No. 8). Then came the idea, 'This stranger might be even worse,' but not very insistently, for, as I looked, I thought, 'Well, this might be the "vanilla,"' and an agreeable feeling tone spread over me."

S. IV.

"I saw both bottles (Nos. 3 and 5) at a glance, and recognised them quickly in turn with brief, faint olfactory images. I knew that I liked both but I liked the green one (No. 5) better."

S. V.

"I wanted to smell the left hand bottle (No. 7) very much, but I did not know it, and I thought, 'It is more reasonable to choose the pleasant odor that one knows.'"

S. VI.

"I recognised both (Nos. 1 and 7) as bad immediately, though I did not remember their odors at all. But while the yellowish one (No. 7) was neutral, I had a strong feeling of dislike towards the other."

S. VII.

"I tried hard to bring back an olfactory image of each, but could only remember them in thought terms as, disinfectant (No. 7) and drug (No. 2). There was an annoyance at not being able to recall their images, but the notion that the yellow was better than the red (No. 2) guided me."

S. VIII.

"I accepted the green (No. 5) at once, since I knew it was pleasant though I could not bring up an olfactory image of its odor. The brown (No. 8) I knew as definitely unpleasant."

One other motive, that of self-assertion, persisted not only as the confirmatory factor in choice, but, occurring once in this capacity, persisted as a penumbra that affected later decisions. Having rejected a certain bottle once, in some cases there was a tendency to reject it on further occasions, in order "to be consistent with oneself." This occurred, not merely where reversals were asked for, but in a series where nothing like such a "prejudice" could be assumed to be aroused by opposition. Its frequency was not great however, being limited to three subjects and occurring with only one sample

from each. This infrequency would be suspicious if the olfactory values could be recognised, but, as they occurred in the first series of presentations, this cause is impossible. The only other plausible explanation is, that the effect of colour associations persisted, but there is no evidence for this.

S. II.

"Did not know either, but having rejected the yellow (No. 7) before, I felt I had to be consistent here."

S. IV.

"No reasons for the choice came at first since I could not recognise either. Then the thought of previous decisions came to my mind. I had always rejected the brown (No. 8), and the thought came 'I'll reject that now and be consistent.' I had no liking for the purple (No. 4), but had no such objection to that."

S. V.

"Remembered rejecting the red (No. 2) before, but had not experienced the odor of either of them. Then the thought came, 'I'll reject the red and avoid changing my mind.'"

It may be inquired as to what is the function of the self-assertive tendency in decision, if it does not always occur as motive. The role that it plays should be evident from the foregoing. Its part is regulative and confirmatory. If motives are not present, as in the case of the Indifference types, then the tendency is very weak; where the other motives are strong, and the verdict is clear, it confirms and strengthens the choice by enhancing the object through a process of intensification of attention, and adds a complacent glow of satisfaction with the choice, leaving a general tone of self-congratulation. This consistency results in a unity of experience that must go far to lift the conduct that it regulates, above the level of pure association and habit, and provide an active source of resistance against change. In most of these effects it is identical with Ach's "determining tendency."

The process of getting experience divides readily into three characteristic periods. The first of these is exploratory in character, and the type of motivation is chiefly associative or simply inquisitive. While the forms of decision include many Indifference or Conflict types, the second period may be characterised as the learning period, and the motive of curiosity—the desire to fix and allocate the relation value—is an

important factor, for, at this period, "unpleasant" odors, whose definite qualities are unknown, yet whose disagreeable side is realised, may be chosen in place of a known agreeable type. Olfactory images are more frequent here than in any other period. The process of choice consumes much time chiefly on account of prolonged efforts to recall past experiences. The final period approaches that mechanised phase investigated by Barrett (3). Here the values have become static owing to past experiences, and the final criteria of choice, viz. purpose values, which are olfactory in the present case, form the sole consideration. The reaction time at this stage becomes reduced to a minimum, recognition of value coming immediately with identification or perception. Choices of the Conflict type occur less frequently in the final phases, and, the novelty having worn off, the Indifference type sometimes recurs here but not frequently. One noticeable feature is the slightly longer time required for choices where intermediate values in the scale are offered. The investigation was not continued to such lengths as completely to automatise these; this point may be corroborated by a reference to distribution of errors in experience.



#### IV. STATISTICAL ANALYSIS

Proceeding upon the principles of analysis indicated in the previous section, an attempt was made to differentiate the responses and tabulate them under the heads of the types that appeared. Naturally, borderline cases were found that had to be arbitrarily placed under one or other forms, but, in the main, there was only moderate difficulty in assigning them to their places. No response was rejected even in the first series, for it was felt that, under certain circumstances, the subject, under practical conditions, would be compelled to force the issue and arrive at some decision even though it should be an indifferent one.

The distribution of the various types is shown in Table II. In the following, and other tables, each type will be indicated by its capital initial letter.

TABLE II.

FREQUENCY DISTRIBUTION OF THE VARIOUS TYPES IN THE MAIN PERIOD.  
Series I.

Types.	P.	C.	I.	J.	Total.
Av.	15.0	7.25	3.75	4	30
A. D.	3.5	1.7	2.2	—	—
% of each	57.7	27.9	14.4	—	100

Series II.

Av.	34.6	7.3	6.1	—	48
A. D.	3.9	4.5	3.9	—	—
% of each	72.1	15.2	12.7	—	100

While the average percentages of the I types approximate one another in each case, there is a noticeable difference in the P and C types. The truer representation of normal cases would, probably, be that found in the second series rather than in the first since, in the first, the questions were of such a type as to arbitrarily include a possibility of numerous conflicts for the purpose of examination of process. On the other hand, the number of these decisions in the second series would better indicate that this approximates an average sampling in the world of affairs, for this process of decision was carried on only till the automatised stage, such as Barrett investigated, was reached, when decision is formulated as rapidly as the alternatives are realised in perception or thought, and the pro-

cess thus automatised. How far this process may be considered voluntary rather than one of mechanical "habit" can best be shown by a comparison of the times and the distribution of the types of decision according to each round of eight experiences.

Fig. 7 shows the growth of experience and the subsequent lessening of times with increased knowledge of values. In the case of certain subjects there is an increase of the initial times due to the effort to recall certain indistinct criteria from previous experiences, S III being the extreme case in this respect. After the fourth round there is a noteworthy general decrease which becomes still more marked at the sixth, where the final range for all subjects drops down to the limits of from one to two seconds. Here is the point of mechanisation, when all extraneous factors, such as curiosity, are eliminated, and the actual criterion is that of final values for the act in question.

Fig. 8 shows the total distribution of types with each round, thus offering means of comparison as experience develops. The C type is noticeably strong in the initial round owing to the total absence of any criteria beyond those of extraneous association. The I type, too, is fairly strong here, this circumstance probably being due to the new experience. The perceptible rise in the fourth round is hard to account for; a comparison with Fig. 7 shows a marked increase of times at this point. There is a possibility that the task had lost somewhat of its freshness, at the same time it had become perplexing in that the values of the odors were not readily recalled from the appearance of the bottles. The absence of clear and definite criteria would thus account for both protraction of times and the number of the I types found, since at this point each subject tended to make an especial effort to recollect olfactory values, the failure in the attempt to fixate them tending to indifference.

In considering the reversals, it is important to observe whether the numbers of types are similar to those occurring in the process of decision, since it was only in approximately fifty per cent of the cases that it was required of the subjects. Table III shows the comparative frequency of these types in percentage form when reversal was prescribed.

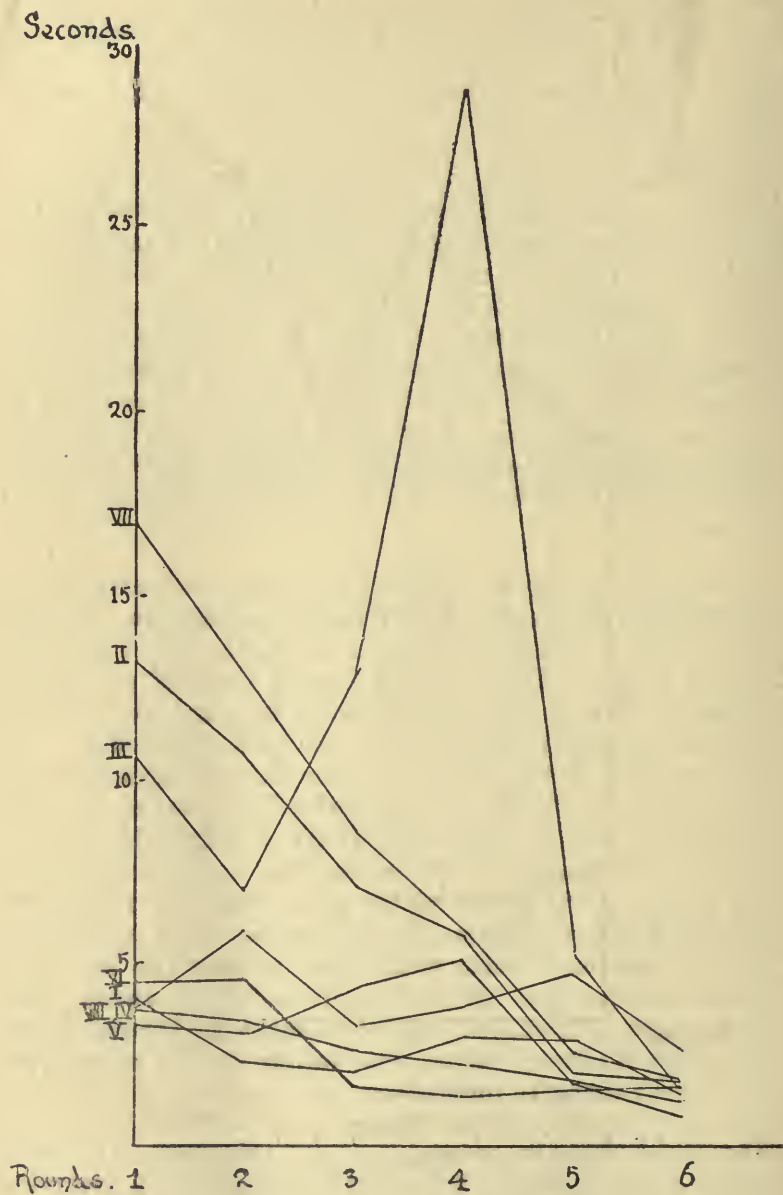


Fig. 7

Average Times for Subject for Each Round of Series II.



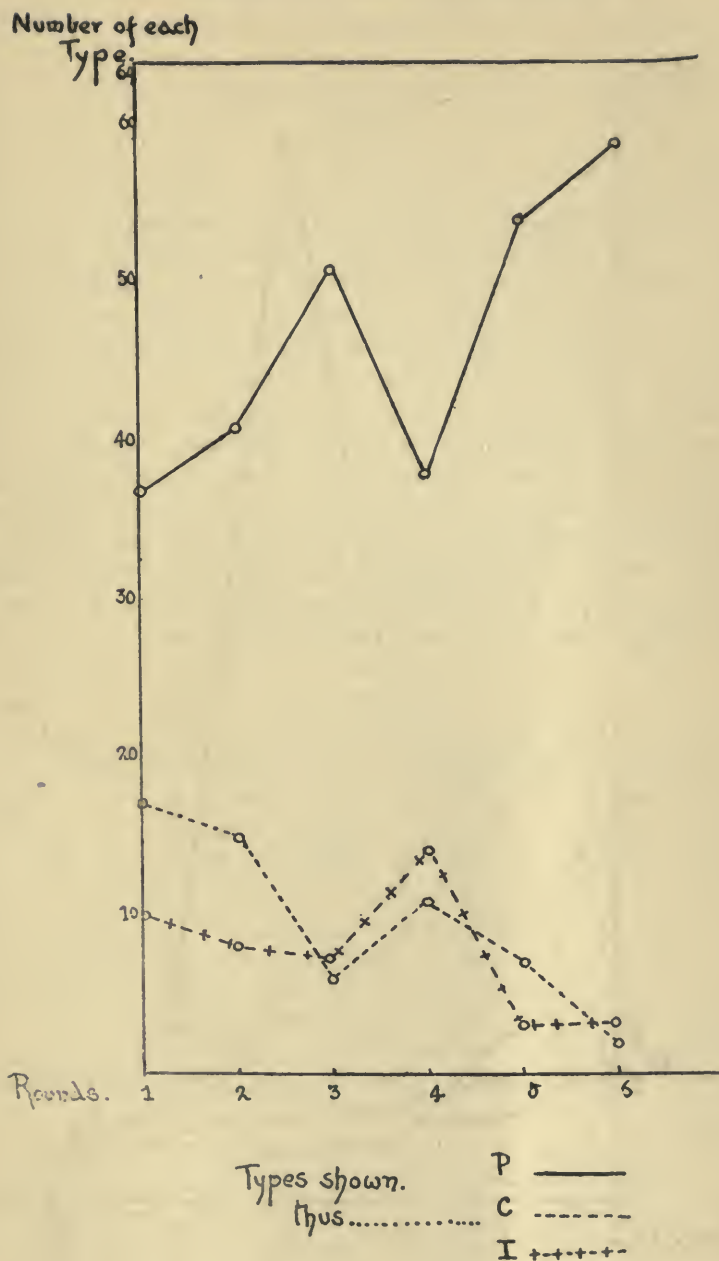


Fig. 8

Graph of Distribution of Types in Each Round of Series II.

TABLE III.

Series I.	P. 52%	C. 33%	I. 15%
Series II.	66.5%	18%	15%

If these results are compared with the percentages of types of the original decisions shown in Table II, the average of the differences of percentage between each corresponding member in each series will be found to be only 4.4 per cent with an A. D. of 2.1, the smallest deviation occurring in the I types. This, in a sampling within the small range of numbers of the examples, would appear to be a safe margin of percentage to allow for a fair estimate as to possibilities of reversal. The actual results of attempts at reversals are shown in Table IV given individually for each series. In each type the term Non-R indicates that the original decision has been reaffirmed, while, in the cases designated R the original decision was reversed and the other alternative accepted. On account of great individual variations complete tables are shown.

TABLE IV.

FREQUENCIES OF REVERSAL OR NON-REVERSAL FOR EACH TYPE.

Series I.		P.		C.		I.		J.		Totals.	
Subjects		Non-R.	R.	Non-R.	R.	Non-R.	R.	Non-R.	R.	Non-R.	R.
I		5	0	3	1	0	3	2	1	10	5
II		2	3	3	3	1	1	1	1	7	8
III		9	0	3	0	1	0	3	0	16	0
IV		5	1	2	1	2	1	3	0	12	3
V		8	0	5	0	0	0	3	0	16	0
VI		7	0	2	3	1	0	1	1	11	4
VII		6	1	3	0	0	2	2	1	11	4
VIII		5	0	4	0	1	2	2	1	12	3
% of each Type.		90.3	9.7	75.8	24.2	40	60	77.3	22.7	77.9	22.1
Series II.											
I		12	2	1	1	1	8	-	-	14	11
II		5	9	1	8	0	2	-	-	6	19
III		17	2	1	2	1	2	-	-	19	6
IV		15	0	6	1	1	2	-	-	22	3
V		20	0	4	0	0	0	-	-	24	0
VI		15	6	0	3	0	2	-	-	15	11
VII		11	1	3	2	0	7	-	-	14	10
VIII		17	1	2	1	0	4	-	-	19	6
% of each Type.		84.2	15.8	50	50	10	90	-	-	66.8	33.2

In each case the noteworthy characteristic is the low comparative percentage of reversals in the P type of decision. Where the evidence is clear and motivation towards one or other alternative decided, then the possibility of reversal appears to be the least. Where the motives are more equally balanced, the possibility of reversal is very much greater in both series, yet not nearly so great as in the I type which presents the most marked tendency in this respect. Though the general trend of results is to increase the possibility of reversal from the P through the C to the I type, one discrepancy between the two results is marked, that is, the increased percentage of reversals in the second series. This may be well explained by the difference of the two forms of situation used. In the former the matter was momentous to the individual, even though hypothesized, rather than actual, and in most cases the individual felt it to be a matter of importance to adhere to the decision. The preponderant motive was often derived from some phase of self-assertion and this original measure of that tendency being ultimately reinforced by the final moment of choice, the result developed into an almost "reversal-proof" type. In the second series the motives were naturally of a simpler nature and of less weight. As a consequence, the strength of the submissive tendency was able to prevail over the motive and its "confirmatory moment" of self-assertion. One other matter is that of the difference between types in the possibility of reversal. This may be regarded as due to the variability of the ratio between the degree of self-assertion and the degree of strength of the submissive tendency in the acceptance of the task. In each case the co-efficient has an individual value and its result is expressed in this direction by the number of reversals achieved.

In considering the times, the fore period of the first series apparently shows no prognostic signs by means of which the type may be forecast; there is no consistent variation in this period by which one experimental type may be distinguished from another as regards duration, not only in the case of the group but more particularly in the case of an individual comparison. Thus, while S. I may show a more prolonged time for the fore period of the P type, S. III and S. IV show exactly the same result for the I type, S. VI and S. VII on the other hand consume the greatest amount of time in the fore periods of the C type. This is more or less in conformity with the



probabilities of the case, for there is no guarantee that the alternatives posited in the main periods will bear any relation as to difficulty with the fore period. The average times of the fore period of the first series are given in Table V while none of course were taken for the second series. The Average Deviations shown are obtained by using the individual averages in each case and not the original single times.

TABLE V.

GENERAL AVERAGES OF TIMES FOR THE FORE PERIOD, SERIES I.

	P.	C.	I.	J.	Av.
Av.	12.90	15.61	11.01	11.86	13.87
A. D.	3.56	5.63	5.58	6.48	3.75

The times of the main period are, on the other hand, distinctly symptomatic in relation to the functions of each type. The individual averages are shown in detail in Table VI, and it is not difficult to see that not only do the general averages agree in relation to the types, but that individual times also bear out this relation. Thus, in each case, taking the general individual averages as a basis of comparison, it will be found that all the P types in both series consistently occupy less time than the averages mentioned, apparently indicating that the decision has been arrived at with a minimum degree of difficulty. On the other hand, the C type exhibits a directly opposite tendency; the times are prolonged by the inability to decide, and the results show a consistently positive tendency for the subject to exceed the average time occupied in his commitment to a final issue. In the case of the I type, no positive tendency in one direction or the other is discernible, the trend being in both positive and negative directions from the average. This again may be explained by the difference between the individual degree of acceptance of the task.

Thus S. II, who exhibited what may be described as "conscientiousness" or sense of duty to a marked degree, shown by the number of reversals achieved, (cf. Table IV) occupied considerably over the average time for the main periods in both series, in achieving these reversals, while S. III, who consistently exhibits a tendency to reaffirm his original decision, from which, together with his introspections, may be deduced a fair degree of self-assertiveness, occupies far less time than his average. Others show a moderate variation in either direction in one series or the other.

TABLE VI.

INDIVIDUAL AND TOTAL TIME AVERAGES FOR THE MAIN PERIOD GIVEN IN SECONDS

Series I.					
S	P.	C.	I.	J.	Av.
I	17.27	23.64	35.62	18.45	23.32
II	18.41	36.13	43.36	9.85	26.33
III	81.34	113.85	72.00	48.60	94.52
IV	17.17	29.05	20.13	14.10	20.85
V	11.66	20.04	—	16.67	14.56
VI	21.12	86.74	14.40	18.55	38.12
VII	19.62	55.20	53.41	40.88	38.58
VIII	31.62	43.05	21.35	47.72	34.59
Gen. Av.	23.32	50.96	37.18	26.85	36.33

Series II.				
S	P.	C.	I.	Av.
I	2.14	3.40	3.45	2.71
II	5.66	8.53	11.70	6.37
III	10.03	56.05	9.82	13.82
IV	2.42	9.98	7.35	4.25
V	2.35	6.92	4.50	3.20
VI	2.30	7.80	2.80	2.79
VII	6.06	14.72	7.28	8.16
VIII	2.22	5.72	3.20	2.62
Gen. Av.	4.14	14.14	6.24	5.49

In the Reversal period, the general average times of both series are shown in Table VII.

TABLE VII

INDIVIDUAL AND TOTAL TIME AVERAGES FOR THE REVERSAL PERIOD GIVEN IN SECONDS.

Series I.					
S	P	C	I	J	Av.
I	29.80	24.13	13.07	25.00	25.69
II	47.52	52.10	28.30	7.45	41.45
III	58.50	86.07	73.20	71.40	66.51
IV	31.42	34.80	18.40	18.93	26.92
V	12.78	14.48	—	9.90	12.89
VI	47.59	78.62	76.60	48.20	56.16
VII	27.93	55.07	27.20	38.73	35.42
VIII	48.93	55.27	33.33	37.30	40.55
Gen. Av.	38.06	53.82	38.58	32.18	38.19

Series II.				
S	P	C	I	Av.
I	5.13	8.60	4.23	5.08
II	31.88	12.71	23.40	24.30
III	10.68	23.40	5.20	11.59
IV	4.01	5.46	6.03	4.66
V	5.99	7.55	—	6.25
VI	7.47	8.27	3.30	7.24
VII	4.02	6.64	4.34	4.66
VIII	4.01	3.73	3.10	3.82
Gen. Av.	9.15	9.54	7.08	8.45

They exhibit a marked increase over the corresponding averages of the main period. The reason for this increased time is evident when referred to the discussion in the previous section. Here the introspective data pointed first to a marked unwillingness to return to a reconsideration of the alternatives; there was a distinct mental set, or disposition, to be overcome, and then, when the subject had succeeded in this, there was often an inability to reconsider the rejected alternative in anything like a favourable light, more especially in the P and the C types. The greatest difficulty appeared evidently in a reconsideration of the P types, for those are the times that show the greatest average increase, and in most cases, are considerably above the average times. On the other hand, the average times for the I types appear to be approximately equal to the main period times of that type, while the C types show a decrease. With the former the same degree of indifference evidently comes into play, while in the case of the C types one of two things occurs; either the original alternative shows an increase, thus metamorphosing it somewhat into the nature of the P type, or the rejected alternative is newly reinforced by the acceptance of the task set in reversal, thus giving a preponderance to the other side, and making possible a reversal of the decision.

An analysis of the number of subjects showing an average increase or decrease of reversal times over the main period, is given in Table VIII.

TABLE VIII

NUMBER OF SUBJECTS GIVING INCREASE OR DECREASE (-) OF REVERSAL TIMES OVER DECISION TIMES.

Series I.	P	C	I	J	Av.
	7	4	3	4	5
-	1	4	4	4	3
Series II.	7	4	3	-	6
-	1	4	4	-	2

In both series there is a uniform prolongation of time in the P type, one subject, S. III being the exception, but it may be observed that his times are abnormally long in the main series. In the C type, increases and decreases are balanced, while the I type appears to have a slight advantage in ease of reversal if judged by the time alone. The analysis of the



increase or decrease in Reversal time is carried out still further in Table IX into R and Non-R decisions.

TABLE IX

GENERAL AVERAGE OF INCREASE OR DECREASE IN REVERSAL TIME OVER MAIN PERIOD TIME.

Series I.					
	P	C	I	J	Av.
Non-R					
Gen. Av.	13.9	8.2	15.0	10.9	10.7
A.D.	7.8	11.2	18.8	26.4	5.4
R					
Gen. Av.	9.2	17.9	-10.4	-1.7	.1
A.D.	4.7	9.5	10.8	14.1	2.7
Series II.					
Non-R					
Gen. Av.	4.9	-.5	1.9		3.9
A.D.	6.2	4.1	6.2		4.7
R					
Gen. Av.	10.1	-3.8	-.2		1.9
A.D.	8.5	8.6	1.4		3.3

As an indication that the attempt was attended with real difficulty it is significant that the Non-R series generally show a positive and greater increase over the reversals. Accepting the time indicator as a significant index of attention, this result seems to point to the fact that the reversal takes place, not so much from the concentration of attention upon the decision, as from the fact that the decision is the factor that determines attention, for the longer period does not result in a change of decision. An examination of the non-reversal series shows that, with only one exception (the C type, Series II), the general tendency was to take a longer time to attempt reversal than to arrive at the original decision. Especially is this the case with the P type which, in both series, stands well above the total averages. The least increases are shown in the C type, but it is to be remembered that, in their case, the original decisions occupied a prolonged period well above the average. Turning to the decisions that were reversed, the P type here shows a consistent tendency to increase the time over the original; in the second series this is particularly noticeable. In the C types the same feature is found in the R series as in the Non-R i.e., that while in Series I this type occupied a far longer time than the original, in Series II the time is far less; the same result is found in degree of difficulty, and may be ascribed to the comparative unimportance of the second series. The I types in both series show a

decrease in time, indicative of the proportionally greater strength of the task over the original decision.

A consideration of the difficulty encountered shows the same trend as the foregoing. In the consideration of both difficulty and confidence expressed numerically, it may be remarked that the rating scale method of treatment is followed, and the significance of the figures in such an application imply merely perceptible degrees of difficulty or confidence, while the final averages can be taken simply to indicate a general trend in one direction or the other. The various average degrees of difficulty in the main period are shown in Table X.

TABLE X  
AVERAGES OF DEGREES OF DIFFICULTY FOR EACH TYPE OF DECISION IN EACH SERIES.

Series I.					
	P	C	I	J	Av.
Av.	1.55	2.73	1.99	1.66	1.95
A.D.	.16	.24	.49	.28	.19
Series II.					
Av.	1.21	2.28	1.65		1.59
A.D.	.14	.20	.45		.28

There is least difficulty encountered in the P type which, in both cases, was well below the general average, while in the C type the highest degree of difficulty was found. On the other hand, the I type approximates the average. This agrees with what was found in the main period in regard to the times, so that the difficulty was evidently associative in character as well as pertaining to the final moment of decision. In the Reversal period of the first series no degree of difficulty was asked for, hence no comparison is possible, but in the second series it was required. The average degrees of difficulty show a fluctuation accordingly as the act of reversal is accomplished or the original choice reaffirmed.

Table XI shows these differences arranged in regard to whether the reversal was effected or not.

Except in the P types the number of cases was not large; the result is fairly large A.D.'s in the C and I types, hence their figures must be regarded as probable rather than final. The apparent contradictory dimensions of the A.D.'s, in comparison with the averages, may be explained by the fact that

TABLE XI

AVERAGE INCREASE OF DEGREE OF DIFFICULTY FOR EACH TYPE IN SERIES II.

Series II.				
	P	C	I	Av.
Non-R. Av.	.75	-.13	-1.00	.45
A.D.	.43	.34	1.00	.23
R. Av.	1.42	-.50	.05	.28
A.D.	.40	.63	.41	.37

the numbers run from minus quantities below zero to positive quantities above zero. Considering each in order, the P type was found much more difficult in attempted reversals than any other, the degree of difficulty greatly exceeding the average in both cases. The reversals were evidently much harder of accomplishment for the subject than were the reaffirmations, for where the reasons of choice were well grounded the effort to reject and replace them would naturally be great, the effort to reconsider would be difficult, and the final reaffirmation would naturally come much more easily than a reversal of decision. In the C types of the second series reconsideration was not quite so difficult as in the first, especially in reversal. The struggle between motives may have been lessened in their case by a reconsideration, the C decision then passing into the nature of a P type or, on the other hand, the previously rejected motive was considerably strengthened by the demand of the task that it be accepted. While it was easier to reaffirm the original decision in the I types, the reversal experience was only about equally as difficult to accomplish as the original decision.

One further measure remains, that of confidence, that is to say the degree of assurance that the subject experienced in connection with his response after having come to his decision. In the first series, the original decision was rated in terms of confidence, but in the after period, in terms of, "possibility of reversal." As far as the experimenter was able to judge, however, the subjects accepted the latter in such a way that it actually stood for confidence in the second decision. The results of the main period are included in Table XII and show only slight variations between the two series.

In each case the general tendency is for confidence to decrease throughout each type of decision. In the P type the



TABLE XII

GENERAL AVERAGE DEGREE OF CONFIDENCE OF THE MAIN PERIOD FOR EACH TYPE OF EACH SERIES.

Series I.					
	P.	C.	I.	J.	Av.
Gen. Av.	3.71	2.45	1.98	3.22	3.12
A. D.	.25	.31	.19	.61	.36
Series II.					
	P.	C.	I.		Av.
Gen. Av.	3.88	2.84	1.89		3.45
A. D.	.09	.21	.23		.22

confidence is strong, approaching in the average far towards a degree of perfect confidence expressed by the number 4.00; the C type fall between a state of being "confident" to a "fairly confident" condition, while, in the I type the range is from "barely confident" to "confident." In the first series this trend is the case with all the subjects with two exceptions, S. III and S. IV whose degree of confidence in the I type is slightly in excess (.11 and .18) respectively of the C type. In the second series, there are no exceptions to type at all, but the average records of all subjects in all cases follow this general trend.

In the treatment of the reversal period the result may best be represented as in the case of degree of difficulty, classified according to increase or decrease in rating on the main period and whether reversed or not. The average group results are shown in Table XIII.

TABLE XIII.

SHOWING GENERAL AVERAGE OF DEGREES OF CONFIDENCE FOR THE REVERSAL PERIOD FOR BOTH SERIES.

Series I.					
	P.	C.	I.	J.	Av.
Non-R. Av.	.28	1.03	.60	.22	.56
A. D.	.19	.16	.33	.14	.27
R. Av.	-1.10	-.21	-.06	-.40	-.67
A. D.	.55	.47	.05	.24	.41
Series II.					
	P.	C.	I.		Av.
Non-R. Av.	.04	1.12	2.00		.23
A. D.	.02	.31	.00		.18
R. Av.	-.65	-.31	.04		-.32
A. D.	.36	.14	.15		.46

Entirely different results appear where the decision was reaffirmed to when it was reversed, and these resemble the findings when the matter of difficulty was discussed. In all cases of non-reversal an enhanced degree of confidence is found, less so in the P type where confidence is already extremely high, but in cases where the confidence is less marked as with the C type, then a reaffirmation results in a greatly increased degree of confidence, strengthening the weaker assurance in this case and bringing it up almost as high as the original confidence of the P type. In the I type, the confidence is also greatly enhanced by reaffirmation, very markedly in the three cases of the second series; the self-assertive tendency evidently took the place of original motive to some extent, or, as sometimes happened, some further motivation developed after the original decision.

If the case where reversals take place be now considered the opposite results are found, a diminution taking the place of an increase of confidence. The greatest decrease is found in the P type. A successful effort to reverse a clearly defined choice appears to be not only exceedingly difficult, but inimical to confidence as well. In the C type the decrease is not so marked but is at least present, while again the I type suffers scarcely any change, the first series of such being slightly negative and the second positive, but the whole range not exceeding from two to three per cent in either direction from the average.

Using the main period as a basis the statistical results of the main and reversal periods may be comparatively summarised in brief tabular form.

In general it will be observed that the P type are principally affected by reversal, and the C type by reaffirmation; the former loses in the essentials of a strong and well formulated decision by change, the latter tends to gain by its own reaffirmation; the changes in the I type are negligible except for the increased confidence in non-reversal. A possible cause of the last effect may be due to the strength of the submissive tendency, which now takes on an opposing direction and provokes the self-assertive tendency by its very opposition. Thus, where a contradictory task is strongly combated and rejected, the result tends to strengthen the original self-assertion or determining tendency. Where, on the other hand, its direction is accepted in opposition to the individual's

TABLE XIV.

SHOWING MAIN CHARACTERISTICS OF TYPES IN THE MAIN AND REVERSAL PERIODS.

## Decision or Main Period.

Type	Frequency	Time	Difficulty	Confidence
P.	Prevailing type	Short	Slight	High
C.	Infrequent	Protracted	Very Pronounced	Moderate
I.	Infrequent	Average	Average	Fairly Low

## Non-Reversal

P.	Frequent	Increased	Increased	Slightly Increased
C.	Evenly Distributed	Unchanged	Unchanged	Greatly Increased
I.	Infrequent	Increased	Decreased	Increased

## Reversal

P.	Rare	Increased	Greatly Increased	Greatly Decreased
C.	Evenly Distributed	Unchanged	Slightly Decreased	Slightly Decreased
I.	Frequent	Slightly Decreased	Unchanged	Unchanged

decision, then its effect is to weaken the strength of decision. In this respect the general findings of the quantitative data bear out the findings of the qualitative analysis of the introspective results.

One final treatment still remains; it concerns the statistical data of judgments. A reference to the tables shows that throughout, these types reach a very close approximation to the general averages. The number of judgments used in this investigation was limited, being only four per subject, and was confined to the first series, so that, while useful as a qualitative check, their paucity does not warrant more than a passing statistical comment. So far as their results suggest anything, they seem to indicate that judgments are similar in general nature to decisions, with a possibility of the occurrence among them of similar types. On this basis they may be regarded as general types, while the decisions themselves are specific and personal solutions of definite situations. Before a final pronouncement could be made, however, a more extended investigation than the present is essential.



## V. DISCUSSION OF RESULTS

It is possible at this stage to gather the findings of the various parts of this investigation, and to compare them with the results of other investigators: the experimental process of choice may be said to begin with the acceptance of the task or Aufgabe in the fore-period. This gives a directing trend to certain associative processes, and develops a general set in the direction of the possible alternatives that may offer themselves. In this regard the process of the fore period may be looked upon as providing a general background for the problem.

In the main period there is a more detailed process of exploration. Here the individual specifically tests the effects of each process by momentarily experiencing first one and then the other of the possible alternatives, by treating them as ideational experiences. If the first attempts are unproductive, the process is continued until something like an evaluation of the motives is arrived at and a comparative weighting is arranged in relation to final purpose. The process at this stage is dependent on past experience either germane to, or bearing directly upon, the situation. This may be termed the associative period. With the completion of this process there is a tendency to accept, deliberately, the alternative regarded as of greater worth by the subject. This acceptance, in contradistinction to the act of judgment, is not merely acquiescent, but active and real, and accompanied by an implied determination of purpose to fulfil the choice in action.

The whole of these processes of tentative trying-out are accompanied by kinaesthetic and organic processes—bodily stresses and strains, which tend to be suspended at the moment of acceptance and to cease with the actual decision. The final course of this motor disturbance ends, not in a complete relaxation, but in a definite kinaesthetic set of preparedness, that Washburn (23, p. 161) has designated in motor terms as a "system of static innervation," its psychological counterpart being a type of "determining tendency." Shifts of attention accompany the tentative acceptances of the alternatives and the final and decisive acceptance of one is attended by its fixation in the focal point of attention.

So far as this description has proceeded it is corroborative of the work of Michotte and Prüm, and of Wheeler, and is only differentiated in the direction of material, the factors concerned in the involved processes being essentially the same. So far as the findings of Wheeler are concerned it differs from them in regard to the cause effecting this final motor set. For him it appears to be associative, working through the mechanical forms of "definitisation, delay, and reinforcement" (24). The present investigation would however, attribute this result to some regulative cause which, for lack of any better or more descriptive term in English, has been designated—following McDougall,—the "self-assertive tendency" and which adds the significance of Ach's phrase, "I really will," to the project. Thus the final act of choice may be said to take on its final significance of a self-determining tendency. There is a subsequent effect of self complacency that tends to mark this process, reported as, "I felt very pleased with my choice," that may be interpreted as a result of this, to-be-gratified tendency.

On the other hand, the "self-consciousness" of Michotte and Prüm is sustained in its reference, upon either content, function, or both. This self consciousness may be conceived of as a regulative tendency as previously described, accompanied by self content. It was shown that self reference was not restricted to content of the process of choice but might be found in certain cases in judgments also, and could not therefore alone be considered as characteristic of the act of decision. The results in this respect agree with Titchener's summing up of the manifestation of the self (21, p. 320): "It need not appear in many of the situations that are designated by self words.....It is the specific expression of a special determination." In this respect its function in the final act of choice is to direct the process at its consummation and carry it over in the form of a definite mental set with a physiological counterpart. Since the arrival at a decision results in a bodily relaxation, the set is probably cortical in character, such as described by Hunter (8), in the higher forms of delayed reaction.

There remains the possibility that this self reference may be analysed further, as Wheeler (24) alleges, into a certain group of elementary and imaginal experiences. That such is the case with the content side can be readily agreed upon;



the imaginative side of purpose, it was found, could be paralleled on the sensory side by organic and kinaesthetic sensations. In so far, however, as the question is one of process or the operation of law, it is impossible to do so. Thus, the chemical composition of sugar may be reduced to the three common elements, carbon, hydrogen and oxygen. But the contiguous presence of these elements even in their necessary proportions does not, on the other hand, constitute "sugar." The substance is something more than an assembling of the three elements; it may be looked upon as a manifestation of a certain process in operation upon these according to the law of such a process. No possible amount of analysis could reduce the process itself to elements. The moral is obvious; while the self content may be reduced to simple imaginal and sensory elements, the vital part of it, that which is really characteristic of the self factor, viz. the self-assertive tendency, cannot be so reduced, but remains unanalysable, an irreducible active principle that combines these elements into a self entity.

When an examination of the various examples was made, it was found that, while the majority of decisions conformed to this particular form, viz. the P or Preference type, there were others that differed in certain respects. Thus in the case of the C types the final decision factor was not so strong as in the case of the former. The same degree of assurance was lacking, and the one alternative was not generally focalised through definitisation and reinforcement. The tendency was for the rejected alternative to persist in a less definite degree than the accepted one, but yet to such an extent as to prevent the whole-hearted acceptance of the latter by the subject. In the case of the third, or I type, there appears to be lacking even an associative basis of self reference, and all that is present is the bare minimum activity of the self-assertive principle. The last have been classed by Michotte and Prüm (1) as "non-voluntary" activities, and result from the bare acceptance of the task in the command to choose. That they are altogether non-voluntary is doubtful however, since there is no further evidence to be considered. The one type that has not been examined in the present investigation is the purely mechanised form, where the relative values have been predetermined, and only the mechanised reaction of decision remains without an actual



weighing of motives. This has been regarded as an habitual, rather than a voluntary activity, and is generally so accepted.

The question that may now be asked is, "How do these experimental types compare with the classification of other writers, more particularly James, since his categories are the most widely inclusive?" The P type may be identified with his "rational" type, and with the predominant type of Calkins and Stout. The designation, "rational" would be better reserved for judgments; the term "preference" as used here appears more truly descriptive of the typical decision process, which is frankly personal, and sometimes, as some of the examples show, almost irrational in tendency. On the other hand the decision with "sense of effort" has much in common with the C type. The conflict here is however, not really ended but persists, as Calkins asserts, in a less intense form. It occurs in practical life, often as the result of motives arising from the self regarding sentiment coming to an issue with those that have to do with the individual's more material welfare, but, so far as this work shows, they are not necessarily confined to these situations, that is to say, the type transcends a mere ethical differentiation of motives, and may be found at any moral levels of conduct; in this respect the C type may be admitted as genuinely psychological. For these reasons, a differentiation between those decisions that include "a moral factor" and those that do not, cannot be regarded as justifiable, and further, such a category violates the logical principles of classification since this type cannot be regarded as exclusive of others, but tends to fall under the heads of Preference or of Conflict. James' second and third classes, where the decision "is given without due deliberation," and where the subject "wearies of the conflict," would probably fall under the I type, unless in the latter case some of them might be included under the head of C. If the subject's experiences are not sufficient to form a basis of experience for the functioning of associative processes, then the result is purely an indifferent one. In the case of the second type, if the subject is not predisposed to either alternative even though his experiences may include both, the decision is unaccompanied by bodily stresses and strains and must be classified as being under the I type. On the other hand, even though the conflict may persist, the subject really makes a deliberate choice; he is well aware that the chosen alternative is his only

reasonable course, and that, though an impulsion towards the acceptance of the rejected alternative may persist, and in an unguarded moment may be followed, it is inconsistent with all his past experiences and his self-assertive tendency in general. Further experiences in such a case tend to render the type a closer approach to the P form.

The relative degrees of strength shown by the self-assertive tendency in these forms is evident when the subject is required to consider a reversal of his choice. Not only is there a change in regard to the difficulty of the task, but there is also a change in the degree of confidence. This is due to the *Aufgabe*, or urge of "oughtness"—originally neutral in regard to the chosen alternative—becoming an active factor working in opposition to the self-assertive tendency or principle of "subjective willingness." The effect on the latter when an attempt to oppose it is made, results in an emotional expression of mild forms of anger experienced as "irritation and annoyance." The result of the new conflict thus aroused depends on the relation between the degree of submission to the task or *Aufgabe*, and the strength of the self-assertive principle. Where the latter is strong, as in the case of the P and C types, the result is a rejection of the new task and a victory for the self. Where, as is generally the case in regard to the I types this relation is reversed, the tendency is towards an acceptance of the *Aufgabe*, and a reversed decision ensues; this is not always the case however, for even the minimum functioning of the self-assertive tendency proved to be too strong in some cases for the reversal to take place.

We may regard the imposition of the *Aufgabe* on the one hand as a manifestation of the submissive tendency—the acceptance of direction from without; it embodies an implicit, "I must comply," a tendency of duty or oughtness. On the other hand the self-assertive tendency carries with it an assertion of determination, or, in the terms of Ach, an "I-really-will" significance. In the first part the subject's acceptance of the task reinforced the determining tendency, submissive acceptance thus supplementing self-assertion at the same point. In the reversal period the two are opposed; the new *Aufgabe* requiring a change in choice, and the still active determining tendency opposing it. Where the choice is vigorous as in the P type, the regulative force of self-assertion is very much

stronger than the tendency to submit, to accept the dictation from outside, and opposition only arouses it more intensely. Where the choice is weak, then the acceptance of the Aufgabe, the operation of the submissive tendency, is relatively the stronger, and the sense of duty overcomes the inclination of the choice.

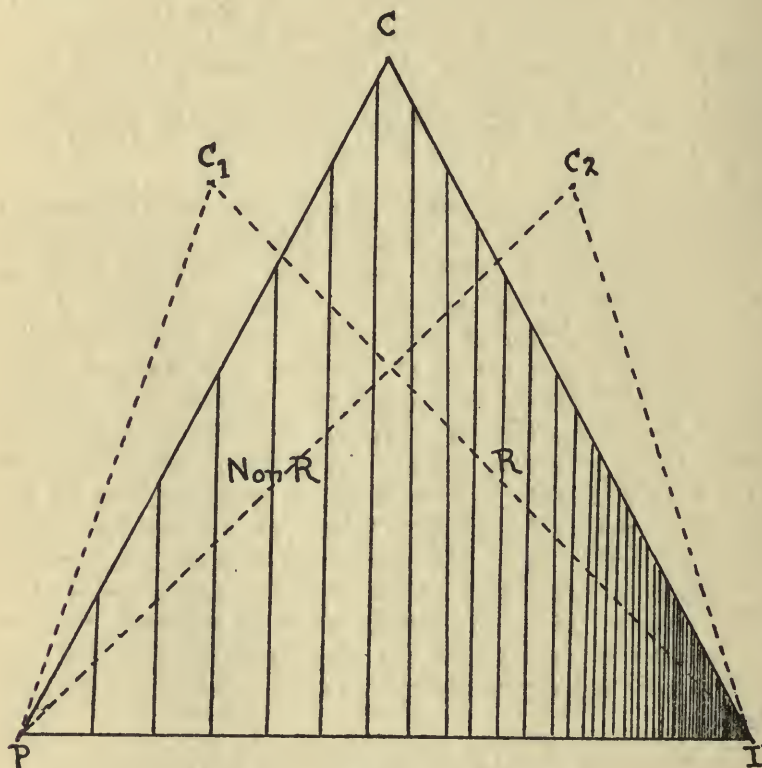


Fig. 9

The relations between the various types and the tendency to merge into opposed forms may be observed in the diagram shown in Fig. 9. Here the three types may be represented in their pure or ideal form by the various angles of the triangle. Where the associative elements are weak, either through lack of the dispositional forces or experience on the part of the subject, these types will fall along the lines, P I and C I. Where they are particularly rich and complex they will tend to fall schematically along the line P C. The line C I would correspond in richness of association with the line P I, with the



angle I marking the zero point. The degree of self-assertive quality would tend to predominate at P with a lessening activity towards the point I. The maximum of possibility of reversal is shown at I in the region of darkest shading. This possibility tends to pass off as the movement is towards P. When the point reaches towards C the chances are evidently somewhere between those of P and I according to other possibilities of the function of the experience and the self-assertive tendency. Given therefore the relative strength of the associative tendency aroused in a decision process, its location between the point I and the line, P C, might be found as a line P' C' parallel to P C.

On the other hand the degree of difficulty may be regarded as the relative approach to equality between the evaluations of the alternatives of choice, reaching a maximum point at C and lessening towards the points P and I. A line P' I' between C and the base P I and parallel to the latter will then be the location of the degree of difficulty. The point of intersection between the line of associative strength, P' C' and the line denoting degree of difficulty P' I' would locate the relation of the decision point to the three types. By dropping a perpendicular to the base line P I, and noting its relative position between the points P and I, the point so found would correspond to the degree of confidence in the decision. If, however, the degree of confidence was found, as in the present investigation, the process might be reversed, and the strength of associative difficulty deduced from the intersection of the perpendicular marking confidence and the parallel line marking difficulty. An appeal to introspective evidence, since no quantitative data were collected in this regard from subjects, tends to confirm the genuineness of this relation.

The relation of the point C to the ends of the base P and I is not fixed equidistantly from each as represented, but its position tends to vary with the individual subject. Where the degree of self-assertive tendency is high C would tend to move towards I, but where it is low the tendency would be for it to move towards the end P, the confident subject tending to reduce the C and I types to a minimum, while these are greatest in number among the less self-assertive types. The number of reversals also corresponds with this trend. Outside the triangle beyond P, would fall all those experiences which have become mechanised in the form of habit, while

in the corresponding position beyond I would fall all those where lack of experience, or individual trend of disposition placed them beyond the possibility of even the semblance of a decision process.

The whole trend of this investigation has been in the direction of removing the emphasis that is usually placed upon primary motives and putting it upon the "actual moment" of decision, viz. the self-assertive tendency. Through the activity of this principle it is possible, even when two favourable alternatives are considered, to choose one completely and fully and to totally exclude the other. If this process of choosing were merely one of competition of alternatives as motives, then the Conflict type with its lack of finality, rather than the Preference type, should predominate, since the subject is initially inclined to both sides. This is not the case however, the Preference type is the predominant and characteristic type, and the final set effected by such a decision leaves but the one alternative clear and distinct, with the other withdrawn from the central field of attention. By the operation of the added factor of self-assertion in a perfect type of decision, the original motive is thus lifted out of the level of competition, and stands alone and focalised as the one purpose of consciousness, and is further maintained there against opposition. Hence, the concern of decision is not with native tendencies, *per se*, but rather their direction and regulation through the work of the self-assertive tendency.

Further there are some significant implications that follow from the conflict that ensues when the self-assertive tendency is brought into opposition with the submissive tendency in the new task. Having made a choice, and called up in support of it the self-assertive tendency, the subject is then asked to change; the endeavour to subdue the original self-assertive tendency and the acceptance of the new alternative constituting the conflict. These implications affect various branches of psychology: primarily an application may be made to the field of psychiatry. Up to the present, reports of such phenomena are confined to individual case reports alone. The subjects furnishing such reports are, from the pathological nature of their condition, more or less suggestible, and totally unacquainted with the nature of psychological processes, and further, their abnormal condition would also tend to affect the veracity of the report. By utilizing and developing such

a form of investigation as indicated here and bringing it under laboratory conditions, such phenomena may be reduced to exact experimental forms, and directly utilise the reports of skilled psychologists. The attainment of both qualitative and quantitative results of importance should follow, and "academic" psychologists would be in a position to offer direct help to the psychiatrist, just as the biologist does to the general practitioner of medicine through the field of bacteriology. Again, in regard to educational and moral training two important principles emerge; these are, (a), that the normal method of procedure should be to endeavour first, to induce the subject to accept the task and then to will it as his own single purpose, (b), after arousing the self-assertive tendency in a given direction not to attempt to interfere by imposing a directly contrary task. The former principle of harmonization of duty and purpose also substantiates psychologically what is ethically postulated as "inner freedom." Finally, the whole of the foregoing has a direct bearing on the laws of mental hygiene.



## VI. SUMMARY

- (1). The process of choice invariably involves a self reference in its purpose; such is not apparent in the case of judgments.
- (2). The final factor in decision is the self-assertive principle, the latter being regulative rather than contributing to content.
- (3). Three differentiated types of decision are apparent, viz. the Preference, the Conflict, and the Indifference type, the first named being characteristic and most numerous.
- (4). The Preference type proceeds smoothly, implies a rich subjective experience and is finally accompanied by a large measure of the self-assertive tendency.
- (5). The Conflict type, vacillating in its character, also implies a rich fund of associations, but appears somewhat lacking in regard to the degree of self-assertion that is present.
- (6). The Indifference type is wanting in associative material, its process is apathetic, and it ranks lowest in the strength of the final self-assertive tendency.
- (7). The possibility of a reversal, (all things being equal), is least in the case of the Preference type, and greatest in the Indifference type.
- (8). Reversal with the Preference type is only attained with great difficulty, occupying a more protracted time than the original and exhibiting a distinct weakening of the self-assertive tendency; in the case of the Conflict type there is a slight decrease in difficulty, time and confidence; with regard to the Indifference type both time and confidence remain unchanged, while the difficulty appears to be moderately increased.

- (9). In these cases where the original decision is reaffirmed, there is a marked increase of time and difficulty in the Preference type, with a slight increase of confidence; in the Conflict type the time is slightly decreased, the difficulty unchanged, but the confidence somewhat increased; in the Indifference type, time and confidence appear unchanged, while the difficulty is slightly decreased.
- (10). The principle of "conflict" set up by the effects of the self-assertive tendency of decision, acting in opposition to the submissive tendency to accept and will a new task contradictory to such decision, offers possibilities of experimental development in the field of abnormal psychology.

## VII. BIBLIOGRAPHY

1. Ach N., *Über die Willenstätigkeit und das Denken*, Göttingen, 1905.
2. Ach N., *Über den Willensakt und das Temperament*, Leipzig, 1910.
3. Barrett E. B., *Motive Force and Motivation Tracts*, London, 1911.
4. Book W. F., *The Psychology of Skill*, University of Montana Publications in Psychology, 1908.
- 5a. Bridges J. W., *An Experimental Study of Decision Types and their Mental Correlates*, Psychol. Monographs, 1914-17, (No. 72) Pp. 72.
5. Calkins M. W., *A First Book in Psychology*, New York, 1912.
6. Gates A. I., *Recitation as a Factor in Memorising*, No. 40, Archives of Psychol., New York, 1917.
7. Hill C. M., *On Choice*, Am. Jour. of Psychol., IX, 1897-8, 587-590.
8. Hunter W. S., *Delayed Reaction in Children and Animals*, Behavior Monographs II, 1913.
9. James W., *Principles of Psychology*, II, New York, 1892.
10. Külpe O., *Introduction to Psychology*, (Trans. Pillsbury and Titchener) New York, 1902.
11. McDougall W., *An Introduction to Social Psychology*, London, 1910.
12. Meumann E., *Intelligenz und Wille*, Leipzig, 1913.
13. Michotte A. et Prüm E., *Etude expérimentale sur le choix volontaire et ses antécédents immédiats* Arch. de Psychol, X, 1910, 113-321.
14. Morgan J. B., *The Overcoming of Distraction and other Resistances* No. 35, Archives of Psychol., 1916.
15. Ogden R. M., *An Introduction to General Psychology*, New York, 1916.
16. Pillsbury W. B., *Fundamentals of Psychology*, New York, 1920.
17. Rowe E. C., *Voluntary Movement*, Am. Jour. of Psychol., XXI, 1910, 513-562.
18. Selz O., *Willensakt und Temperament. Eine Erwiderung*, Zsch. f. Psychol., 1911, 59, 113-122.
19. Stout G. F., *Manual of Psychology*, London, 1913.
20. Titchener E. B., *A Text Book of Psychology*, New York, 1910.
21. Titchener E. B., *A Beginner's Psychology*, New York, 1916.
22. Warren H. C., *Human Psychology*, Boston, 1920.
23. Washburn M. F., *Movement and Mental Imagery*, Boston, 1916.
24. Wheeler R. H., *An Experimental Investigation of the Process of Choosing*, No. 2, University of Oregon Publications, 1920.
25. Woodworth R. S., *The Cause of a Voluntary Movement*, Studies in Philosophy and Psychology, Commemorative Volume, Boston, 1906.
26. Wundt W., *Outlines of Psychology*, (Trans. Judd C. H.), Leipzig, 1907.



# SOME WELL-KNOWN MENTAL TESTS EVALUATED AND COMPARED

BY  
DOROTHY RUTH MORGENTHAU, Ph. D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

*No. 52*

NEW YORK  
MAY, 1922

226736<sup>2</sup>  
7. 11. 28

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Carey St., W. C.); Paris (16 rue de Condé)



## TABLE OF CONTENTS

Introduction	5
Subjects	8
Tests Briefly Described and Reasons for Their Selection	14
Method—Applying Tests to Subjects	21
General Considerations	
Specific Observations on the Application of the Test Selected	
Results	25
Conclusion	52



### ACKNOWLEDGMENT.

For the advice of Professor Edward Lee Thorndike of Teachers College, Columbia University, and of Dr. William Healy and Dr. Augusta Bronner of Judge Baker Foundation, Boston, the writer wishes to express appreciation. Special thanks are due for the painstaking assistance given by Professor Robert S. Woodworth of Columbia University, New York City.

## Some Well-Known Mental Tests Evaluated and Compared

ONE who approaches the subject of the measuring of children's mentality will find that the mind of the normal child has received attention in what we may call vertical and parallel respects. There have been a considerable number of tests developed by students of psychology in the endeavor to secure mental measurements independent of the experience and judgment of the clinician. The development has been in a vertical manner, that is, the best recognized psychologists who have undertaken this work have each developed tests, have each put them into extensive practice and have published the results of that experience. But each of these psychologists has developed his test on his own suppositions, and, basing the nature of his test on his own experience, has tried to evolve a plan of testing which is supposed to be useful in determining mental conditions of such general extent that they may roughly be called intelligence. Thus we have the Stanford-Binet scale, the most generally used of any one of the mental tests. Then there are the Porteus tests, the Pintner-Patterson performance scale, and a dozen or more of others which are known to every clinical psychologist.

The development of mental tests has been parallel in that none of these psychologists in developing their own ideas have carried them to the point of thoroughly comparing the results obtained by their tests to the results obtained by the simultaneous use of a number of the other tests all with respect to normal children. There has been some comparison of results of the various tests when applied to abnormal children but this has not been thoroughgoing and has been done not by making the tests with the idea of eventually combining the results and of placing valuations upon them, but merely in the course of clinical work with abnormal children. It is questionable whether such results are sufficiently thorough to be considered the basis for a convincing answer as to the relative value of the respective tests, and inasmuch as they were made on abnormal minds, one would not dare to trust even those

comparative results with respect to what the test will show as to normal minds.

Those who have developed their respective tests have compared them with *some* other mental test, most frequently one of the Binet revisions. But, no considerable number of the tests which have been so developed in parallel fashion have been applied purposely to obtain comparative results and to ascertain which if any of them can be shown to be untrustworthy and what group of them can be relied upon as furnishing a satisfactory schedule for testing and comparing the common elements of mentality in normal children.

Upon perceiving that there was a lack of any purposely made comparative study of mental tests it was proposed herein to set forth the results of such a study of about a dozen of the most commonly used mental tests. The tests were applied to a large number of unselected normal children, in general each child receiving the full schedule of tests. By means of the results to be obtained from this comparative study it was anticipated:

1. That the degree of reliability of each test would be indicated.
2. That the same purposes could be effected with respect to the value of each test.
3. That the information obtained under the first and second headings would make it possible to select a schedule of tests of indicated reliability for application to normal minds, or further, whether the Stanford-Binet alone would suffice.
4. That by restricting the ages of children tested in general to from ten to sixteen years, the period in which individual capacities first assume importance for vocational determination, it would be possible to guide the vocational training with some degree of success:

A brief statement of the results can now be given reserving the more detailed statement involving the basis and methods for the results for future pages. The first aim, to secure an estimate of the reliability of the tests used, was largely successful. Of the thirteen tests, the reliability of which was investigated, one class, the four construction tests, Healy A and B. and Knox Moron tests, and diamond shaped frame, were found to be unreliable; five other tests were found to be reliable, namely the Stanford-Binet, Pintner Non-Language group test, Thorndike Reading Scale Alpha 2, Porteus Maze test, and



Tapping test; while the reliability of four tests, the Myers Mental Measure, Healy Pictorial Completion test II, Healy-Bronner learning tests, and the Crossline test was undetermined for various reasons.

The results obtained as to value of the tests were as follows:

Stanford-Binet, Pintner, Alpha II, and Porteus are valuable tests and should be included in individual case studies. In spite of their unmeasured reliability, Myers and Pictorial Completion II are also valuable tests and should likewise be included. Judgment should be suspended with regard to learning tests. The Tapping test is of doubtful value and its use should be left to the discretion of the examiner. The Construction tests because of their unreliable character do not give valuable results.

As to the schedule of tests to be used in testing normal minds it was found best not to use the Stanford-Binet alone but to have the schedule composed of that test and the five others which were found valuable. From the tests used and results obtained it cannot be stated here whether this schedule is of value as to vocational guidance for the reason that the factors involved in each test are not known with certainty and until they are known, definite valid conclusions about the abilities of the individuals concerned cannot be reached.

## SUBJECTS.

It was desired to test one hundred normal but otherwise unselected children. In order to obtain an unselected group it proved necessary to select the subjects very carefully, for, if all the children tested had been from a Children's Home, or from a Settlement, or from any one school, the result would have been a highly selected group. To avoid this a few were taken from many different sources and in this respect the distribution proved to be reasonably satisfactory.

As to age, originally the plan was to have about ten children at each of the ten periods of one year each, from seven to sixteen inclusive. But this plan was given up because our interest is not with the six or seven year old who has to go to school and learn fundamentals, no matter wherein he is gifted and who rarely shows talents or handicaps at such an early age. Our chief concern is with children in the sixth, seventh or eighth grades and in high school, because they are the adjustment problems, and because it is important to aid them if possible in deciding whether they should remain in school or go to work. If the latter what should they do, if the former what sort of training do they need? So the attempt was made to lay all the emphasis here and reduce the number of children under eleven to a minimum. Another objection to the original plan is that ten in a group is too small for any kind of generalization.

The total number tested was 128, of which 116 usable records were retained. For various reasons many of these records are incomplete so that this number was necessary in order to have a minimum of 100 scores on each test. There still remain some tests which were given to less than 100 children, but the number is in each case sufficiently large to give valuable results.

All defectives were excluded, for in mixing their records with those of normal children many confusions would have arisen, and the issues would have been less clear. Much intensive work has been done in testing defectives, so that we know a great deal about their reactions to a group of tests such as we have chosen. To be sure, they vary considerably in their results, but we know in general the points where they are weakest as in abstract reasoning and formal generalization,

and also the points in which proportionately they excel. By narrowing the field to normals the significance of the conclusions can be made more pertinent. This was an arbitrary procedure dependent largely upon the judgment of the writer, and subject to criticism on this basis. It is quite possible that some very dull normals were also excluded, this being justified on the grounds that their normality might reasonably have been called in question by more severe examiners. With reference to the three cases whose I. Q.'s fall below 80, there seems to be no doubt that they are to be considered as dull normals. The grade they attained in school for their age, their response on the other tests and their behavior in the community all argue for including them in our study. The boy receiving the lowest I. Q.—73—was born in the United States but taken to Italy at the age of five, and remained there six years. In spite of this he was in the eighth grade. He did very well with all the construction tests.

As no limitations were set at the other end, the grade and I. Q. distributions are higher than one would otherwise expect in a general sampling of the population.

I. Thirty-seven children, twelve girls and twenty-five boys, were tested at the Home for Jewish Children in Dorchester, Massachusetts. Many of these children were half orphans, some had lost both parents—most of them were in the Home temporarily. They were chosen from the total number entirely by chance. They all attended public school in the vicinity and all but two or three had come to the Home within two years. All were able to speak and understand English, this being the only language used at the institution, although in many of their homes no English was spoken. Their ages ranged from 7-0 to 15-1.

II. Twenty-four girls came from Frances Willard Settlement in Boston, Massachusetts. These were divided into three clubs—one consisting of one seventh grade and ten eighth grade girls, the youngest being 12-7 and the oldest 14-2. They came one evening a week for the express purpose of taking the tests. They were the first ones to volunteer from a large group. The other two groups of seven and six respectively were younger girls who happened to meet on afternoons which were convenient for the examiner.

III. Six high school girls in New York volunteered to take the tests.



IV. The ninth grade consisting of six boys and five girls in the Woodmere School (private) at Woodmere, Long Island, were tested. The ages ranged from 13-1 to 15-2.

V. The poorer section of the 8B class of Public School 11, New York, were tested. There were thirty boys in the class ranging in age from 13-2 to 16-11.

VI. Finally eight miscellaneous children were tested.

The subjects selected appeared to give a satisfactory difference in quality so as to bring out the capacities of the tests to meet a variety of normal mental conditions.

TABLE I AGE DISTRIBUTION 116 CASES

Yrs.	Mos.	Frequency	Yrs.	Mos.	Frequency	Yrs.	Mos.	Frequency
7	0	1	10	4	2	13	8	0
7	1	0	10	5	1	13	9	2
7	2	0	10	6	0	13	10	3
7	3	0	10	7	0	13	11	5 19
7	4	0	10	8	0	14	0	1
7	5	0	10	9	0	14	1	1
7	6	1	10	10	0	14	2	2
7	7	0	10	11	1 6	14	3	1
7	8	0	11	0	1	14	4	2
7	9	0	11	1	0	14	5	2
7	10	1	11	2	0	14	6	2
7	11	0 3	11	3	2	14	7	3
8	0	0	11	4	1	14	8	0
8	1	0	11	5	1	14	9	0
8	2	0	11	6	2	14	10	1
8	3	0	11	7	0	14	11	6 21
8	4	0	11	8	0	15	0	1
8	5	0	11	9	0	15	1	2
8	6	0	11	10	1	15	2	5
8	7	0	11	11	2 10	15	3	1
8	8	0	12	0	3	15	4	0
8	9	0	12	1	1	15	5	2
8	10	1	12	2	1	15	6	2
8	11	1 2	12	3	2	15	7	2
9	0	1	12	4	0	15	8	0
9	1	0	12	5	2	15	9	0
9	2	1	12	6	0	15	10	0
9	3	0	12	7	2	15	11	1 16
9	4	1	12	8	1	16	0	3
9	5	0	12	9	2	16	1	0
9	6	0	12	10	4	16	2	1
9	7	2	12	11	3 21	16	3	1
9	8	1	13	0	1	16	4	1
9	9	1	13	1	2	16	5	0
9	10	1	13	2	1	16	6	1
9	11	0 8	13	3	0	16	7	1
10	0	0	13	4	0	16	8	1
10	1	0	13	5	3	16	9	0
10	2	0	13	6	1	16	10	0
10	3	2	13	7	1	16	11	1 10
		15			44			57

Distribution of the subjects by age.

It will be noted that only 19 of the 116 subjects are under eleven years old.

TABLE II  
GRADE DISTRIBUTION—114 CASES.

Grade	Frequency	2 had left school.
I	1	
II	2	
III	2	
IV	11	
V	7	
VI	16	
VII	10	
VIII	44	
IX or I H. S.	12	
X or II H. S.	4	
XI or III H. S.	5	
XII or IV H. S.	0	
Left School		
VIII	1	
II H. S.	1	

The vast majority of subjects were in the VIth to IXth grades inclusive.

INTELLIGENCE QUOTIENT DISTRIBUTION. 112 CASES.

Scale:—1 square to 1 child

70 means 70.000 to 79.999 etc.

The curve of distribution is skewed positively.

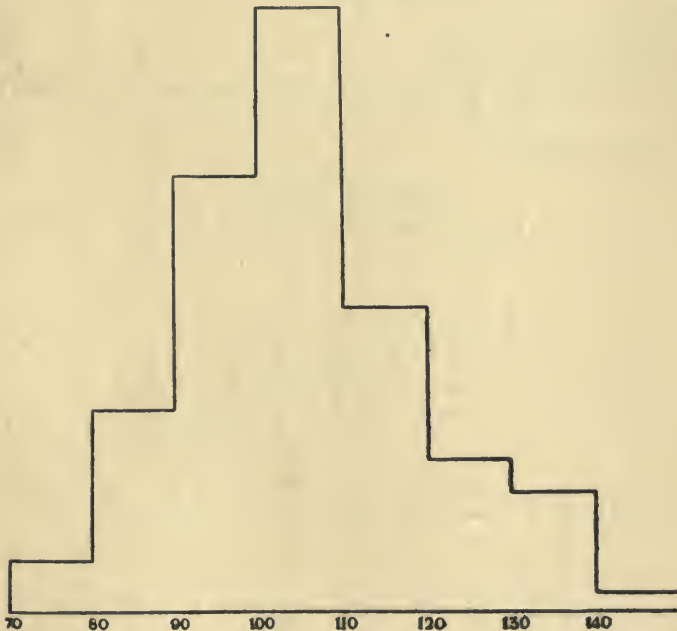


TABLE III

DISTRIBUTION OF INTELLIGENCE QUOTIENTS—112 CASES.

I. Q.	Frequency	I. Q.	Frequency	I. Q.	Frequency
70	0	95	2	120	0
71	0	96	3	121	1
72	0	97	4	122	2
73	1	98	2	123	0
74	1	99	3	124	0
75	0	100	7	125	1
76	1	101	7	126	2
77	0	102	4	127	2
78	0	103	2	128	1
79	0	104	0	129	0
80	1	105	5	130	3
81	0	106	4	131	1
82	0	107	3	132	1
83	0	108	2	133	0
84	1	109	2	134	1
85	2	110	1	135	0
86	2	111	1	136	1
87	1	112	1	137	0
88	2	113	3	138	0
89	3	114	2	139	0
90	1	115	1	140	0
91	2	116	4	141	1
92	1	117	1	142	0
93	4	118	2	143	0
94	4	119	2	144	0

4 were not given the Stanford-Binet test.

Average 104.5

Mental age in months

Average 154.8

Mean Square Deviation 34.54

The table shows that very few of the children tested had I. Q.'s below normal.

#### Age-Grade Distribution

CHRONOLOGICAL AGE												Total cases
	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5	
I		1										1
II		2										2
III			1									1
IV			1	6	3	2						12
V				2	3	1	1					7
VI						6	10					16
VII						1	5	2	2			10
VIII							5	9	15	11	6	46
IX								7	3	1		11
X								1	2	2		5
XI										1	4	5
Total	3	3	2	8	6	10	21	19	22	15	10	116



## CHRONOLOGICAL AGE

Chronological age—mental age distribution.

	6.5	7.5	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5	Total cases
6.5												1
7.5		1										1
8.5		2		1	1	1						5
9.5				3	1	1	1					6
10.5			1	2	3	2	2					10
11.5			1			1	3				1	6
12.5				1	1	3	4		3		1	15
13.5							4	1	4	1		10
14.5						2	3	2	6	6	3	22
15.5								2	4	8		14
16.5							3	4	2	1	1	11
17.5								2	1		1	4
18.5								3	1		1	5
19.5								1			2	3
Total		3	2	7	6	10	20	17	20	16	10	112

Mental age—grade distribution.

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	Total cases
6.5												
7.5	1											1
8.5		2		1	1	1						5
9.5				3	2	2						7
10.5			1	6	1	1	1					10
11.5				1	1	2	2	1				7
12.5					2	4	2	6				14
13.5						3	2	5				10
14.5						4		15	2		1	22
15.5								11	1	1		13
17.5								1	1	1	1	4
16.5							1	4	5	1		11
18.5								1	2		1	4
19.5									1		2	3
Total	1	2	1	11	7	17	8	44	12	3	5	83

## TESTS BRIEFLY DESCRIBED AND REASONS FOR THEIR SELECTION

The large number of tests available had to be classified so as to find which tests covered identical ground; only one of these was then selected. Time was an element particularly to be recorded since preferably less than three hours should be devoted to each child for the completion of all tests. This allotment of time is considered by most authorities to be generous, particularly since the Stanford-Binet takes nearly three quarters of an hour, thus leaving only two hours for all the other tests. Consequently between alternate tests apparently serving the same purpose the briefer one was chosen. The same limitation on the amount of time to be spent on any one individual caused the necessary omission of some tests which were highly desirable except as to their length. In the last mentioned class are group tests requiring an hour or more to be applied. Where the results that are sought can be reached by group tests doubtless much time can be saved in using them, but the inquiries involved herein were such as to necessitate largely individual testing.

In selecting the tests another danger that was realized and that it was attempted to avoid, although, as the results show, not with entire success, was that a great many tests involved many sides of mental activity so that the final result expressed numerically would not be indicative of which mental abilities had tested favorably and which unfavorably. For instance, ability to deal with abstract and with concrete material may be extreme opposites giving a correlation of minus 100. If both kinds of material are combined in one test, the child who succeeds in one may fail in the other and vice versa. In computing the final scores compensation will give the same net result to two children of exactly opposite capabilities. If general intelligence is what we want we may find it in this way, but if we are interested in special abilities or disabilities these tests which hide them must not be used. We have found this confusion to exist in many tests, of course, never in such an extreme form as in the illustration above, and undoubtedly introduced on purpose, but we feel that its value is at all times questionable. This error is extremely difficult to eliminate

completely, in fact we can not be sure even now, as it will appear later in the results, that we have successfully done so.

Another source of error too often overlooked was borne in mind in selection of the tests, namely the variability of the test that is being considered. Where the same test is applied to a person at intervals and it is found that the resulting scores are not identical the question arises whether the varying scores can be combined so as to give a reliable standard for use and comparison with the results obtained when the test is applied to other children, or whether the variation indicates an unreliability in the test itself sufficiently serious to warrant the test being discarded. As an example of variations of such minor character that their existence does not indicate unreliability, and which can be compensated, we can take the tapping test where there may be a variation of about five taps in each direction from the average, which would be entirely satisfactory. Such variations are due to unessential and insignificant details of the conditions under which the test is repeated, such as posture of the child being tested, kind of pencil or stylus being used, etc. Taking ten or more measures of tapping ability would increase the reliability but the final results would show such slight difference from the result of one or two trials that the frequent repetition is entirely uncalled for to secure reasonable reliability.

On the other hand, if the variations in result obtained by repeated use of a test on the same individual are not of a minor character and if the day-to-day variability is so erratic that the variation is all the way from good performance to poor performance, then the situation is either that the child tested is shown to be subject to mental disturbance, or that the test itself shows a high and dangerous variability. If it is the test that is variable, it is obviously essential to weed it out *ab initio*. Such variability has been found to exist in the Knox cube test, in the application of which a uniformly normal child may make the record of an imbecile one day and of a super-normal child the next day. Of course, such a test, if not eliminated, would lead to results that are valueless for comparative purposes and dangerous for diagnostic ones.

As to variability, the reliability of a number of tests was established and recorded before the study was undertaken. As to the remaining tests, in order to overcome the possible existence of variations indicating unreliability it was necessary to



retest each child with the same or with a similar test after an interval of a week—no less or practice effect would be met, no more to avoid the effect of any mental growth in the interval.

The necessity of retesting caused by possible variability in the test itself, led to the subordinate but difficult problem of determining what methods of retesting would avoid errors due to the process itself. Thus, as has been mentioned, retesting must be done in such a manner as to avoid practice effect. It has been shown by various workers that certain types of tests once solved, such as most puzzles, are no longer tests at all, whereas others, such as auditory memory for digits and psychomotor control, show a minimum effect, which, after the week between tests, is negligible. Those of our tests which come within the last-mentioned class were similarly repeated. Those which were of the former type had similar tests substituted for them in the second trial, while still others falling between these classes were altered in details so that the same test could be repeated, avoiding the memory aspect.

The tests finally selected were:

1. *The Stanford revision of the Binet-Simon scale.* This test is so widely known that it does not seem to be necessary to describe it here.

2. *Pintner's mental survey non-language group test*, with Myers Mental Measure as an alternate for repeating. These tests involved a minimum use of language. In the Pintner test no language is used in the performance, and in fact it is possible to give this test to foreigners or deaf children through the medium of signs, while in giving the Myers Mental Measure it is necessary for the subject to understand simple language, but none is used in executing the test. The Pintner test has six parts, the first resembling the Knox cube test, the second and third being substitution tests, the fourth a drawing completion, while the fifth is a reversed drawing test, and the sixth a picture reconstruction. Following directions, Pictorial Completion, and two tests of picking out objects with common elements, compose the Myers test.

3. *Thorndike's reading scale Alpha 2.* This is a test in which language plays a prominent part. The subject reads a paragraph and then reads certain questions based upon the paragraph to which he writes his answer. To succeed he must understand the context of the paragraph, he must understand the question and know what it calls for, and he must be able

to find the answer in the context and write it down. This is a graded test which is applicable from the second grade through high school. Since the practical work of this research was undertaken, Dr. McCall of Teachers College has considerably increased the usefulness of this test by devising ten sets identical in method but with different contents, of which the test here used is one. It is now known as the Thorndike-McCall reading scale and its reliability has been thoroughly established.

4. *Healy's Pictorial Completion Test B* is an apperception test with the language element omitted. The ten pictures (plus one sample) present a day's activities of a young school boy, in which each picture contains a situation known to every child, such as eating breakfast, the school cloak room, a street accident, etc. In each picture one important element is lacking; pieces which complete the picture, plus fifty more of the same size being arranged in a definite order in a box from which the subject is at liberty to choose those which he desires. A clue to the missing piece is furnished by the pictures.

5. *Porteus Maze Tests*. Vineland Revision 1919. These tests are supposed to measure social fitness and common sense. Among the capacities which they were devised to measure are forethought and planning capacity, prudence and mental alertness in meeting a situation new to experience. There are eleven mazes, graded in difficulty from year three to fourteen. Beginning with year five, avoidance of blind alleys is the main requirement for a successful performance. The more complex the maze, the further ahead must one look in order to be certain that one is choosing the correct path. There is no time limit; in fact no mention of speed is made, and if the child asks he is told to do it as well as possible, taking as long as he likes. Porteus says that children fail mainly because of impulsiveness in action, overconfidence and carelessness, lack of pre-consideration, lack of planning capacity, irresolution and mental confusion, inability to sustain attention, or to profit by past mistakes.

6. *Tapping Tests—Healy's Form*. This consists of a sheet containing one hundred and fifty half inch squares, arranged ten in a row—fifteen rows. The subject taps once in each square, without touching the lines and covers as much ground as he can in thirty seconds. This is a simple test of psychomotor control which was repeated without alteration.



This test in a slightly different form was first introduced by Cattell in 1896, for testing freshmen at college. He had one hundred 1 cm. squares, into each of which the student must put a dot, completing the task as quickly as possible. Time was recorded; evidently there were no errors. This test was supposed to measure rate of movement. Clark Wissler used it with many of Cattell's other tests in his "Correlation of Mental and Physical Tests" on college freshmen in 1901. He found that the average time for men was 34 seconds, for women 30.8 seconds. In 1911 Whitley: (M. T. Whitley, *An Empirical Study of Certain Tests for Individual Differences*) reports results on Cattell's test, in which she kept the time constant (30 seconds) but computed the length of time which it would take to complete the blank. We have found the additional fifty squares useful in that some of our cases marked over one hundred squares in the thirty second time limit.

7. *Healy's Construction Tests A. and B.* The Knox-Moron test and Knox Modification of Healy A—a diamond-shaped frame, were used as alternates. We have called these A and B respectively to correspond with the Healy tests and for convenience. The equipment for these tests consists of a board containing one or more openings into which the child tested is supposed to fit pieces of wood so shaped that when properly arranged they will just close up the apertures. An advantage of these tests is the convenient size of the materials required. As all materials had to be carried from place to place the use of clumsy form boards or the tapping board with its dry batteries, metal plate and stylus, was practically out of the question. Where other things were equal, tests having the least paraphernalia were to be preferred.

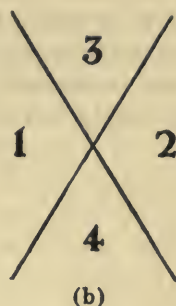
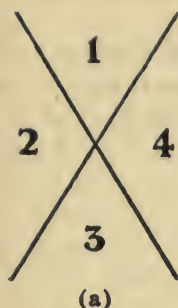
8. *The Crossline Tests* shown in the figure were also given.

The crossline tests were included because they are a modification of the famous Code test, which is generally considered one of the best in the whole Stanford-Binet series. They take very little time to give and can easily be modified for repetition.

9. *Healy and Bronner Learning Tests.*—These tests were devised to test learning ability, not as in the skill experiment, but as it is found essential in the elementary school subjects. Learning test A—the association of two symbols, a figure and a number, resembles other substitution tests such as those of Woodworth and Wells, Pintner, and especially Woolley. The



## I. Crossline Test



## II. Crossline Test

1	4	7
2	5	8
3	6	9

(c)

1	2	3
4	5	6
7	8	9

(d)

(a) and (c) are the forms used generally.  
 (b) and (d) were used for retesting.

difference lies in the fact that three trials were given and speed of learning determined success. Learning test B is the association of a symbol with a sound, as in learning a language. The symbols are from the Phoenician alphabet, and the sounds consist of one or two consonants and a vowel, simple enough to pronounce but without meaning. This prevents older children from forming associations which would be impossible for those who did not know the meaning of the syllables. Test C is the association of a symbol and a value presented audibly, and test D is the association of ideas with a picture. The first three test a sort of rote ability whereas the latter tests learning of ideas.

It seems reasonable that success in school work may depend as largely upon learning ability as upon mental capacity, especially in the early grades where the chief requirement in most of our schools is a good rote memory, as in learning multiplication tables, and these two do not necessarily go together. Certain clinical cases bear out this suggestion, and these tests

were included to ascertain the reactions of normal unselected children in this respect.

*National Intelligence tests* were not yet published in November and December, 1919, when this study was begun, or they would surely have been considered and very likely used.

## METHOD—APPLYING TESTS TO SUBJECTS

### GENERAL CONSIDERATIONS

All of the tests except the non-language group tests and the Thorndike Alpha 2 were given to the subjects individually. The non-language tests were given sometimes individually and sometimes in groups of about ten with one exception where thirty eighth grade boys were tested in a group.

The time of day at which the tests were given varied considerably. About fifty of the subjects, from the Children's Home and from the Settlement, were tested in the evening. All others were tested in the daytime.

Care was taken to avoid giving any tests while the subject might be fatigued. Each child was questioned regarding the matter and whenever there were indications of fatigue the testing was always postponed.

Usually a subject was tested for only an hour and a half at one time; frequently the duration of the testing was shorter and only occasionally was it longer.

The tests were all scored according to the directions laid down by their respective authors. They were all scored personally by the examiner twice. In all of the tests selected for use the scoring is objective and requires no technique. Where possible, score cards or keys were used. Where the time taken by a subject to complete a test was to be recorded, the timing was done by means of a stop watch.

Much effort was expended in persuading the subjects to give an equal amount of attention and concentration to all of the tests, so that the results would not be affected by individual preferences. For a large proportion of the subjects the incentive of vocational guidance was offered and some general vocational advice, based partly on the experience of the examiner as well as on the tests, was given at the conclusion of the testing. Younger children needed no incentive and their enthusiasm was so pronounced that they continually applied to take more than the regular number of tests.

Supplementary information concerning the subjects was gathered and recorded, especially age in months, school grade, success in school work, marks, standing in class, whether a



repeater and how often, whether subject skipped any grades, etc. The vocational plans and interests of the older children were obtained whenever they had any. Results of physical examination were obtainable for a large per cent of the cases. Several subjects had also been given neurological examinations. Occasionally some result can be explained by reference to these findings, as for instance an unaccountably poor performance on the Healy Pictorial Completion test which was probably due to uncorrected vision. One case where peculiar results were obtained from the tests was explained by the physical examination which showed a history of epilepsy and thereupon the case was no longer considered.

#### SPECIFIC OBSERVATIONS ON THE APPLICATION OF THE TESTS SELECTED

*Stanford-Binet.*—In the United States there have been several revisions of the Binet-Simon test, the most recent and well the best of these being that by Professor Lewis M. Terman of Leland Stanford University, California, published in its final form in 1916. This revision, called the Stanford-Binet, was the one used in this study. The score obtained in the Stanford-Binet test is expressed in years and months, mental age. This mental age, when divided by the life age, results in the intelligence quotient, which is expressed as a decimal. There have been some wrongful uses of the intelligence quotient. It is an attractive but erroneous idea that a certain intelligence quotient can be found below which all can be considered feeble minded while all above are normal or supernormal. The error in this idea has been pointed out by Fernald, Mateer, Kohs, and others, who demonstrate the degree of overlapping, and show how valueless the I. Q. is when reported without reference to life age.

The Stanford-Binet results can be analyzed, as well as summed up in the I. Q., and it is possible that a detailed analysis of the data would yield all the information required. The plea that general intelligence scales have a right to be so called is largely based upon the supposition that the functions which are tested are manifold. Auditory memory for rote material and for ideas, visual memory, language ability, reasoning ability, apperceptions, general information and many other abilities—all are found within the total range of tests. Unfor-

tunately, in the Stanford scale no child gets tested in all these fields, and further, since they are not standardized separately the significance of success or failure in one part is difficult to determine.

The Stanford-Binet tests were all given by the writer in the manner described by Terman. It is unnecessary to repeat this test in order to establish its reliability as the reliability has been independently reported upon by Terman.

The vocabulary and memory span for digits of the Stanford-Binet were given with the Porteus tests, the remainder of the Stanford-Binet taking only one session.

The Alpha 2 Reading Scale was scored by the method worked out by Kelley and his tables were used.

The tapping test was scored for number of taps and errors.

In the construction tests number of moves and time were taken and when the test was not completed within the limit of five minutes it was scored as a failure and the number of moves up to that time was noted.

A construction test—once solved—is much easier to solve a second time unless the first solution was due to chance. Healy A was repeated in order to check the first performance. In Healy's construction test B a second trial generally brings a result as near perfect as possible (that is, dependent only on skill and speed in motor performances), even if the first solution was hit upon by chance. It is impossible to do away with the chance element in performance tests, but in order to guard against it as much as possible, two tests were used each time, and the selection was made after a study of many types. There are several difficulties in making this choice and we were impressed by the fact that most performance tests have not been standardized and that there are very few tests of this kind which are sufficiently difficult for older subjects. The Healy and Knox tests satisfied both of these conditions.

The scoring for the learning tests is rather complicated. A perfect score on all four tests is four hundred, one hundred being the perfect score for each test. Learning test A has twelve elements, and if these were all correct on the three trials, thirty-six elements would receive a mark of one hundred, or each would get 2.8. Thus the score equals the number correct multiplied by 2.8. When a perfect score is made on the first or second trial, it is assumed that further trials would give a perfect score also. In learning test B there are five

symbols in each of the three trials,—consequently each receives a value of 6.7. In test C there are seven symbols and three trials. Dividing one hundred by three times seven there results a value of 4.7 for each, while in test D, which has ten items, the total number is thirty, with a value of 3.3 each. The total for all the tests is the sum of the score on each of the four.



## RESULTS

Where a clinician is generally satisfied to take the score obtained by applying a test as a final goal, if in fact he goes so far as to work out a score, it is obvious that to attain the purposes here in mind the scores of the various tests used must be compared to gather statistics reflecting their qualities. That is, when the one hundred and sixteen children had been given the tests that were selected and when the scores were recorded, the field work was completed, but there remained to investigate in a laboratory manner what a combination of the results would show with reference to the purposes of this study.

This comparison of results was made by correlation, that is, by measuring the mutual implications (see Thorndike, *Mental and Social Measurements*, pp. 156-185). A test is to be evaluated in three ways; its correlation with criteria other than results of tests; its self-correlation, and its correlation with other tests. In the present inquiry we obtained no outside criteria with which to correlate our tests, because no outside criteria available could be relied upon. In the field of mental abilities, the only criteria which have been widely used are teachers' opinions, school marks, etc. These are unsatisfactory at best. Although we possess all these data for our cases we consider them useless since the children attended eight different schools in four places, with the marking systems varying for each. We compared judgments as to intelligence made by the teacher of the ninth grade of the Woodmere school with those made by the eighth grade teacher of the New York Public School. In the former the I. Q.'s varied from 95 to 141; in the latter from 73 to 116. In the former all but two children tested as supernormal and the class average was 121, whereas in the latter only one tested above 110 with a class average of 96. But to read the teachers' judgments one would think that the pupils of the latter school were considerably more intelligent than those of the former. Even the comparative ratings within one group were markedly unreliable. They showed all the errors of judgment pointed out by Terman. No account was taken of age; the best behaved, most conscientious pupil was invariably considered the most intelligent, etc. What is the use of making correlations with this kind of material,

when one knows in advance that all the fault of a low correlation will be attributed to the criterion, and the tests will stand as before—unknown quantities! Moreover, these criteria could only be used to represent a measure of general intelligence. The teachers admittedly knew practically nothing about the special abilities of their pupils; the parents, where consulted, knew very little more. A rating on general intelligence has been frequently correlated with general intelligence tests, and the results published. Our data would present no new factors.

Consequently we have not evaluated the tests by means of correlation with outside criteria but we do have the data for self-correlations and for inter-correlations. Where various tests which we used intercorrelate extremely highly, we may feel that they are measuring the same thing. On the other hand, if the intercorrelations approach zero or are negative, the results indicate that we have no evidence that aspects of intelligence are being measured at all. Only if the correlations are sufficiently high to indicate that intelligence is being measured and low enough to show that different factors are entering into the different tests, can we consider the tests worthy of being included in mental examination. In judging our correlations we must remember that we are testing normal children only,—therefore our coefficients are lowered—and that our ages do not cover a large area, which also lowers the coefficients of correlation.

Our conclusions are limited to the tests we used but the general method of dealing with the scores has a wide applicability.

Table IV  
DISTRIBUTION OF PINTNER SCORES—100 CASES

Score	Frequency	Score	Frequency
200-209.9	0	370-379.9	3
210-219.9	0	380-389.9	5
220-229.9	1	390-399.9	2
230-239.9	0	400-409.9	4
240-249.9	0	410-419.9	5
250-259.9	1	420-429.9	5
260-269.9	0	430-439.9	7
270-279.9	1	440-449.9	4
280-289.9	1	450-459.9	5
290-299.9	2	460-469.9	7
300-309.9	1	470-479.9	5
310-319.9	1	480-489.9	1
320-329.9	2	490-499.9	4
330-339.9	4	500-509.9	6
340-349.9	2	510-519.9	4

TABLE IV—CONTINUED

Score	Frequency	Score	Frequency
350-359.9	6	520-529.9	1
360-369.9	8	530-539.9	2

16 were not given the Pintner Test.

The evenness of distribution of scores is noticeable.

Average=420.964.

Unreliability 6.9.

Mean Square Deviation=68.95.

Unreliability 4.9.

Table V

## DISTRIBUTION OF MYERS SCORES—90 CASES

Score	Frequency	Score	Frequency	Score	Frequency
16	0	46	0	76	1
17	1	47	1	77	1
18	0	48	3	78	0
19	0	49	1	79	0
20	1	50	2	80	0
21	0	51	1	81	2
22	1	52	4	82	1
23	1	53	4	83	0
24	0	54	2	84	1
25	1	55	2	85	0
26	1	56	0	86	0
27	1	57	3	87	2
28	2	58	2	88	0
29	2	59	2	89	0
30	0	60	2	90	0
31	1	61	1	91	1
32	0	62	3	92	0
33	2	63	3	93	0
34	2	64	3	94	1
35	0	65	2	95	0
36	3	66	0	96	0
37	1	67	0	97	0
38	1	68	2	98	0
39	0	69	3	99	0
40	1	70	1	100	0
41	1	71	1	101	0
42	0	72	1	102	0
43	4	73	1	103	1
44	0	74	0	104	0
45	2	75	0	105	0

26 were not given test.

Average=53.325

Unreliability 1.8.

Mean Square Deviation=17.88.

Unreliability 1.3.

Table VI

## DISTRIBUTION OF ALPHA SCORES—107 CASES

Score	Frequency	Score	Frequency	Score	Frequency
3.6	2	5.4	1	7.3	7
3.7	0	5.5	1	7.4	6
3.8	0	5.6	1	7.5	12
3.9	0	5.7	1	7.6	2
4.0	0	5.8	0	7.7	6
4.1	3	5.9	2	7.8	1
4.2	1	6.0	0	7.9	2
4.3	0	6.15	1	8.0	2



Table VI—CONTINUED

Score	Frequency	Score	Frequency	Score	Frequency
4.4	0	6.2	3	8.1	1
4.55	1	6.3	0	8.2	2
4.6	0	6.4	2	8.3	2
4.7	3	6.5	1	8.4	1
4.8	1	6.6	3	8.5	2
4.9	2	6.7	5	8.6	0
5.0	1	6.8	4	8.7	0
5.1	3	6.9	3	8.8	1
5.2	6	7.0	1	8.9	0
5.3	0	7.1	0	9.0	1
		7.2	7		

9 were not given The Alpha Test.

Average=6.834.

Unreliability .116.

Mean Square Deviation=1.20. Unreliability .082.

TABLE VII

DISTRIBUTION OF PICTORIAL COMPLETION TEST SCORES—110 CASES

Score	Frequency	Score	Frequency	Score	Frequency
-15 to 0	2	30 to 34.99	6	65 to 69.99	12
0 to +5	2	35 to 39.99	6	70 to 74.99	7
5 to 9.99	2	40 to 44.99	6	75 to 79.99	4
10 to 14.99	1	45 to 49.99	10	80 to 84.99	9
15 to 19.99	4	50 to 54.99	8	85 to 89.99	4
20 to 24.99	4	55 to 59.99	12	90 to 94.99	2
25 to 29.99	0	60 to 64.99	8	95 to 99.99	1

6 were not given test.

Average=54.527.

Unreliability 2.16

Mean Square Deviation=22.69. Unreliability 1.5.

TABLE VIII

LEARNING TESTS DISTRIBUTION—106 CASES

Score	Frequency	Score	Frequency	Score	Frequency
170-179.9	2	250-259.9	4	330-339.9	4
180-189.9	0	260-269.9	7	340-349.9	4
190-199.9	1	270-279.9	6	350-359.9	9
200-209.9	0	280-289.9	5	360-369.9	7
210-219.9	2	290-299.9	8	370-379.9	9
220-229.9	1	300-309.9	10	380-389.9	2
230-239.9	4	310-319.9	8	390-399.9	4
240-249.9	1	320-329.9	7	400-409.9	1

10 were not given the tests; 3 none at all; 7 not all four.

Average=300.66.

Unreliability=5.08.

Mean Square Deviation=52.31. Unreliability=3.6.

TABLE IX

PORTEUS SCORES DISTRIBUTION—113 CASES

Score	Frequency	Score	Frequency	Score	Frequency
5	2	8.5	4	11.5	14
5.5	1	9	2	12	7
6	0	9.5	6	12.5	15
6.5	1	10	6	13	15
7	3	10.5	8	13.5	4
7.5	4	11	11	14	7
8	3				

3 were not give this test.

Average=11.09.

Unreliability .19.

Mean Square Deviation=2.02. Unreliability .13.

TABLE X

DISTRIBUTION OF CROSSLINE TEST SCORES—114 CASES

Score	Frequency	Score	Frequency	Score	Frequency
I II		I II		I II	
Both OK <sup>1</sup>	70	OK <sup>1</sup> -OK <sup>2</sup>	1	OK <sup>1</sup> -F	3
OK <sup>1</sup> -OK <sup>2</sup>	13	OK <sup>1</sup> -OK <sup>3</sup>	2	OK <sup>2</sup> -F	1
OK <sup>2</sup> -OK <sup>1</sup>	5	OK <sup>1</sup> -OK <sup>4</sup>	1	OK <sup>3</sup> -F	2
OK <sup>3</sup> -OK <sup>1</sup>	5	OK <sup>1</sup> -OK <sup>4</sup>	2	OK <sup>4</sup> -F	1
OK <sup>1</sup> -OK <sup>4</sup>	2	OK <sup>2</sup> -OK <sup>4</sup>	1	F -F	5

OK<sup>1</sup>=Correct on first trial.OK<sup>2</sup>=Correct on second trial.

F=Failure on fourth trial.

TABLE XI

DISTRIBUTION OF TAPPING SCORES. AVERAGE OF 2 TRIALS—113 CASES

Score	Frequency	Score	Frequency	Score	Frequency
40 to 44.99	1	65 to 69.99	11	90 to 94.99	5
45 to 49.99	3	70 to 74.99	13	95 to 99.99	5
50 to 54.99	5	75 to 79.99	23	100 to 104.99	2
55 to 59.99	9	80 to 84.99	20	105 to 109.99	0
60 to 64.99	9	85 to 89.99	6	110 to 114.99	1
Average=73.43.			Unreliability 1.26.		
Mean Square Deviation=13.39			Unreliability .89.		

TABLE XII

DISTRIBUTION OF CONSTRUCTION AND KNOX—TIME 108 CASES

Score	Frequency	Score	Frequency	Score	Frequency
50 to 99.99	1	350 to 399.99	12	650 to 699.99	6
100 to 149.99	7	400 to 449.99	8	700 to 749.99	4
150 to 199.99	5	450 to 499.99	8	750 to 799.99	2
200 to 249.99	10	500 to 549.99	9	800 to 849.99	3
250 to 299.99	10	550 to 599.99	5	850 to 899.99	2
300 to 349.99	8	600 to 649.99	7	900 to 949.99	0
				950 to 999.99	1

Average=420 to 480 or 7.685.

Mean Square Deviation=3.39.

Tables 4 to 12 inclusive show the distribution of scores on the various tests. The average, or more properly speaking the arithmetic mean and mean square deviation, are also given for each.

That we have sufficient cases is shown by the relation of the variability to the average. In only a few instances is it large enough to raise a doubt as to whether enough cases were used. These are the Pictorial Completion test, the Construction tests, and the Myers Mental Measure. The formula for the unreli-

ability of an average is  $\sigma T\text{-obt.av.} = \frac{\sigma \text{dis.}}{\sqrt{n}}$  for the unreliability

of a mean square deviation it is  $\sigma T\text{-obt}\sigma = \frac{\sigma \text{dis.}}{\sqrt{2n}}$  These data

are also included in the tables. (See Thorndike, Mental and Social Measurements).

A few special considerations arose at once with reference to the crossline tests, the tapping test and the construction tests.

The crossline test has no value for our subjects (see table X.); one hundred and fourteen cases were tested, out of which 70, or over 60 per cent, made perfect scores; the remaining 40 per cent ranging almost indifferently from one error to complete failure. This test is, then, far too easy for our subjects, and the results are useless for our purposes. We will disregard it completely from now on.

In dealing with the tapping test we were confronted with the problem of how to handle the errors. Since a perfect correlation would be expected between two absolutely perfect tests of tapping ability, the highest correlation obtainable is presumably the one which best accounts for the errors. On this assumption the two trials of fifty cases of the tapping test were correlated both by Pearson and Spearman formulae, first disregarding the errors, then weighting them one each, and finally weighting them two points each, with the following results:

	Pearson	Spearman
Errors disregarded . . . . .	$r = .794$	$r = .917$
Errors weighted one each . . . . .	$r = .773$	$r = .90$
Errors weighted two each . . . . .	$r = .764$	$r = .82$

It would seem then that the errors are of comparatively little importance, but as disregarding them gives the highest self-correlation, they will be omitted in any correlations in which the tapping test is involved.

A similar problem is presented by the construction tests, where we have scores for time and moves: Should they be combined and if so, how? If not, are they both important, or only one, and if the latter, which one? In order to arrive at an unbiased conclusion—for it was the writer's opinion that time was by far the most valuable measure—the advice of fifteen other persons was sought. These others were all familiar with the tests, and had used them extensively in clinical work. By far the majority were in favor of using both time and moves, each independently of the other. Two of these considered the moves decidedly more important than time; two others stated that time alone was sufficient, because time and moves had been found to correlate so highly, that the difference between using them and not doing so was within



the probable error of either one. None recommended attempting to combine them.

The following correlations were therefore made:

Construction A with B-time.

Construction A with B-moves.

Knox A with B-time.

Knox A with B-moves.

Average Construction A and B with average Knox A and B-time.

Average Construction A and B with average Knox A and B-moves.

If the test was not completed in five minutes it was scored as a failure and the number of moves up to that time recorded. Some children who solve the test in three minutes make more moves than others who fail in five minutes. How can one tell how many moves the latter would have made, had they completed the test? Obviously, the number they made until they were arbitrarily stopped is not a fair measure. It was finally decided to omit all cases where any construction test was a failure, from the moves correlations.

The crude scores were not used in the time correlations, but the three hundred seconds were divided into twenty groups of fifteen seconds each. Anyone succeeding with a test in fifteen seconds or less, was put in group one; if he took more than fifteen seconds and less than thirty-one seconds he was put in group two, and so forth. All who failed the test were put in group twenty, thus making it possible to include in these correlations many cases which had to be excluded from the correlations of number of moves made.

Taking up first the self-correlations, that is, the correlations of our alternate tests, with each other or the correlations of the scores obtained by repeated use of the same test, the results were as follows:

As only a few Stanford-Binet's were repeated, the results are of little significance. We obtain a correlation of .89 on our fourteen cases. L. M. Terman ("The Intelligence of School Children" ch. IX.) had retests given to three hundred and fifteen children, out of which forty-six were given three or more tests. The interval between the first and second testing ranged from one day to seven years. The central tendency of change is represented by an increase of 1.7 in

I.Q.; the middle fifty per cent of change lies between the limits of 3.3 decrease and 5.7 increase. Consequently the probable error of a prediction based on the first test is 4.5 points in terms of I.Q. The correlation between all the testings is .933. Apparently whether the interval be a few months or several years does not influence the result. If the re-examination be within a few days, the I. Q. will—on the average—be raised only two or three points, and this when no restriction has been put on the children communicating with one another. There are several exceptions to this general rule, one being that young feeble-minded children tend to show their feeble-mindedness more as they grow older; that is, they test lower on the Stanford-Binet. We need not concern ourselves with this, as only normal children were included in this study. Another obvious factor which tends to make the I. Q. appear unstable, is due to the fact that the test is limited at the upper end. As a child with a high I. Q. grows older, the I. Q. drops until at the age of sixteen years the highest I. Q. obtainable is 122. In many pathological cases such as children suffering from epilepsy, chorea, etc., the I. Q. fluctuates considerably. But even within the ranges of normality, Terman thinks that fluctuations occur for at least three reasons.

1. There may be a certain amount of irregularity in the actual rate of mental development.

2. The results of a test may be influenced to some extent by the conditions under which it is given, the state of the child's health, his attitude toward the test, fatigue and other temporary and accidental features.

Retests after a brief interval indicate that errors from this source are ordinarily not large.

3. There is inevitably a certain amount of error in every I. Q. rating due to imperfections in the scale used.

What has been generally criticized in the Stanford-Binet scale, namely that it measures different things at different years and consequently that a subject might do very well when his memory ability for example was tested, and very poorly when his reasoning ability came into the foreground a couple of years later, does not seem to be valid on actual findings. The theoretical argument against such a criticism is that so many age levels are tested each time that a subject will win and lose points in every branch which the test includes.

The Pintner and Myers tests were chosen to measure the same thing, and so we expected to find a high correlation between them. The Pearson coefficient of .584 was so unexpected that we felt that further investigation was needed. A closer study of the tests revealed the fact that their likeness rested on negative similarity; neither involved the use of language, but in other respects they apparently required different abilities. The Pintner test appeared more limited, more mathematical, involving concrete situations rather than generalizations while the Myers on the other hand was more general, but rather sketchy. In order to test the truth of this hypothesis, the six Pintner tests were intercorrelated and also the four Myers tests—see table. The average of the Pintner intercorrelations was .234, of the four Myers tests correlated each with all the others, .445. It will therefore be seen that the above explanation is unsatisfactory.

TABLE XIII  
PINTNER TESTS INTERCORRELATED

	1	2	3	4	5	6	Composite
1		-.009	.392	.325	.183	.337	.618
2	-.009		.470	.022	-.022	.107	.396
3	.392	.470		.361	.224	.316	.757
4	.325	.022	.361		.035	.456	.625
5	.183	-.022	.224	.035		.307	.540
6	.337	.107	.316	.456	.307		.670
Average	.249	.126	.353	.240	.154	.305	
Composite	.618	.396	.757	.625	.540	.670	

Average of all above correlations, regarding signs + .234.

Average of all above correlations, without regarding signs + .238.

Probable Error of each correlation, approximately .05.

Number of cases 100.

The fact that correlations between the separate Tests are low, while those of each Test with the composite of all 6, are high, indicates merit in the Test as a whole.

TABLE XIV  
MYERS MENTAL MEASURE INTERCORRELATIONS

	1	2	3	4	Composite	Average of all above correlations— + .445
1		.470	.469	.564	.786	Number of cases—89
2	.470		.346	.424	.796	
3	.469	.346		.403	.477	
4	.564	.424	.403		.775	
Composite	.786	.796	.477	.775		

The comment made concerning the previous Table—Pintner Tests—applies to some extent to the Myers Test also. However, the correlations between the separate tests are much higher than those found between the Pintner Tests.



For if the Pintner tests were all of the same nature, including the same factors, their intercorrelations would be high. On the other hand, if the Myers tests were general, their inter-correlations would be lowered. Just the opposite occurs; the Pintner intercorrelations are lower than the Myers. These correlations can probably be explained on another basis. In the Pintner series certain tests are easier than others, most especially the second and fourth, which lowers the intercorrelations. In the Myers Mental Measure all the tests with the exception of the third are of about the same difficulty, the grading being within the test, and this raises the correlations. It also seems probable that while the Pintner tests do measure more limited factors, each test may measure a different one, the type of material alone remaining the same.

On a *priori* grounds something of this sort seems likely, for the material is practically the same, the correlations are low, so the factors measured must be different.

It is true that in the Myers Mental Measure the ability to respond to the spoken word (directions) is part of the test, and it is possible that this is a special ability—calling forth something akin to the abilities necessary for success with the Stanford-Binet, even where the language itself is easily understood. Such a factor our data are unable to measure, but it is interesting in this connection to compare the correlation of the Stanford-Binet with Pintner and of the Stanford-Binet with Myers.

In devising the Pintner non-language test, the effort was made to have it extend from the lowest to the highest grades. This meant introducing tests such as the second, which is far too easy for a child after he has reached the fourth or fifth grade, and also others which were almost incomprehensible to the young child, as tests four and six. Since our subjects are for the most part past the fourth or fifth school grade, we would expect to find some sign of their maturity in the correlations. Reference to the table shows that test two correlates lower with all the other tests than any other single test. The one exception is the correlation of tests two and three, which—it will be remembered—are identical in form, the latter being different from the former only in degree of difficulty. Test six, on the other hand, correlates higher with every other test, than test two. This is as it should be: had we tested younger children the table would

probably have shown entirely different results. Incidentally, these findings show the importance of bearing in mind the nature of the group that is being studied when interpreting correlations. Each part was also correlated with the total test score, with high results throughout, with the exception of test 2. In looking at this table, one must feel that the test is a good one, for the intercorrelations of the separate tests are low, but with the composite they are high.

Before leaving the Pintner test, mention should be made of a study by Jeanette Chase Reamer, in which she retested over four hundred children with this test with slightly less than a two-year interval, and found a correlation of .726 between the relative positions which they occupied at each testing. The closeness of this correlation was a complete surprise to both her and to Professor Pintner.

With regard to the Myers Mental Measure intercorrelations, we find them all fairly high and regular. The most surprising thing about them is that tests three and four which appear far more similar than any other two tests in this series, should have one of the lowest correlations,—lower than four with one or four with two. Also, we see no reason why one and four should correlate higher than any other two. If the language factor were significant, we should find one and three (where audible directions must be followed for each separate unit of the test) correlating highly, and also two and four (where after the original directions the subject is left to himself). But as a matter of fact, one and two, and one and four are higher than one and three, and two and four. However the degree of difference between the various correlations is so small that these comparisons must be taken in a negative rather than a positive sense; that is, we might have expected the correlations to prove *something*, instead of which they prove nothing! With the composite the separate tests correlate very highly, as would be expected since the composite includes always the test being correlated with it, thus giving a perfect relation between two out of the five factors. Test three proves an exception here also, and we feel that the fault is the same as with Pintner two: it is too easy for our subjects.

The Porteus Maze Test when correlated with itself gives a correlation of .95, which is high and satisfactory.

The intercorrelations of the construction tests gave the



most disconcerting results of all. They seem to prove Professor Thorndike's assertion that no matter how many construction tests are used, one cannot do away with the chance element. If four construction tests, when correlated for time and moves, give only .16 for the former and .08 for the latter, it seems like a hopeless task to give sufficient tests to raise the correlation to the high 70's or 80's. This is indeed a problem, for the construction test as such is undoubtedly desirable.

Perhaps more important than these low numerical results, is the fact that combining the individual tests does not seem to operate to raise the correlations. Thus the two Healy tests when correlated for time, give a result of .21, the two Knox tests similarly correlated give .27, but the average of Healy tests with average of Knox tests shows a correlation of only .16. In attempting to explain these findings, it must be remembered that the Healy tests were given on one occasion, and the Knox tests at least one week later, both A and B on the same day.

If our results were due primarily to lack of reliability of the construction tests from day to day then scores from two construction tests given on the same day ought to show higher correlations than we found. If the lack of reliability of construction tests from day to day is not to be considered because of the generally low correlations, and so if it makes no difference whether all four tests are given on the same day or different days, then our average correlations should not turn out to be lower than the correlations of tests given on the same day because an increase in the number of factors generally operates to raise the correlations. Other correlations between various combinations of the construction tests were made and are recorded in table 16, but the results are no more enlightening than the ones we have discussed here.

We are assuming here that the solution of each of the four construction tests involves the same abilities, not that they are of equal difficulty. We have no evidence to prove that this is the case, but we do not see how any construction tests could be devised which, though different, were apparently more similar than these. However, Healy test A and Knox test B are more similar than any other combination. Knox's test was modelled directly from Healy's and is supposed to be more difficult. A correlation between these two gives us



minus .055 for time and .126 for moves. In other words, the correlation between the two is about what one would obtain between two factors having no relationship to each other at all. If this is true of Healy Construction test A and the Knox diamond-shaped frame test, we conclude that construction tests have no constant value for intelligence testing.

It will be recalled that the Thorndike Reading Scale, Alpha 2, was not repeated as the alternate scales now available had not yet been published, but we might quote Dr. McCall's statement to the effect that a high correlation was obtained between our scale and the more recently devised alternate on representative subjects.

No correlation was obtained between two trials of the Healy Pictorial Completion test II because the number of cases who were retested was small, about twenty-five, no further retesting being done because the attitude of the subjects was so different on the second testing that the repetition was more a matter of memory than anything else.

In repeating this test a week or more after the first presentation it was found that the correct pieces were again put in. Of those that were incorrect about half were the same, the other half being pictures having about the same value in scoring, so that the total score was very little altered. It was generally slightly increased, rarely lowered. The only other test of this kind available is Healy's Pictorial Completion test A, which is so simple that almost all of our subjects would make perfect scores on it. The attitude of the subjects, when this test was offered a second time, was not good. The test appeals because it is a new situation presenting a problem in an attractive form. The second time, the newness has worn off. The usual response is, "I've done that before," or words to that effect. If the child is urged to attempt a better performance, he will often ask in a surprised tone of voice, "Didn't I do it perfectly before?" Even when one succeeds in getting a child to try again, he rarely makes any effort, but puts in at once the pieces selected before or similar ones. If he comments audibly on his performance, it runs something like this, "Oh, that one,—a book was missing there; where is it? Here—why there are two—well it doesn't matter, it's a book he dropped." Occasionally one will notice that it *does* matter, but even this is due largely to chance, to his happening to have spied two books this time.

As a whole it was felt that what was gained by repeating this test in the way of establishing its reliability, was not equivalent to what was lost in the attitude of the subjects to the tests as a whole. If repeated at the very end of the testing, this difficulty would be in part eliminated, but it was decided to omit its repetition completely.

The reliability of the Healy and Bronner learning tests was not ascertainable as no alternate series has been devised, and as no other test could be found which appeared sufficiently similar to warrant the hypothesis that it measured the same thing.

The tapping test was repeated in exactly the same form, and showed an intercorrelation of .81 with a P. E. of .022. This we may consider a satisfactory correlation, showing that the test has a high degree of reliability.

The self-correlations having been thus completed and analyzed, the next step is to consider the intercorrelations of the tests. Let us now consider the correlation of each test with the Stanford-Binet mental age. As all the tests are given crude scores regardless of age, in order to have comparable data, the mental age must be used instead of the I. Q.

TABLE XV

	Stanford-Binet	(probable error)	Number of Cases
Pintner	.439	$\pm .055$	97
Myers	.686	$\pm .037$	89
Alpha 2	.757	$\pm .027$	106
P. C. II	.541	$\pm .045$	110
Learning	.491	$\pm .049$	105
Porteus	.536	$\pm .045$	110
Tapping	.604	$\pm .040$	112
4 Construction	.426 (Time)	$\pm .078$	107
4 Construction	.326 (moves)	$\pm .097$	83
Healy A	.410 (Time)	$\pm .079$	110
Healy A	.374 (moves)	$\pm .092$	88
Healy B	.281 (Time)	$\pm .088$	110
Healy B	.088 (moves)	$\pm .105$	88
Knox A	.046 (Time)	$\pm .095$	111
Knox A	.009 (moves)	$\pm .099$	103
Knox B	.216 (Time)	$\pm .090$	111
Knox B	.112 (moves)	$\pm .098$	103

Total number of tests correlated with Stanford-Binet  $r = .5976$ . Woodworth's method of combining the results of

several tests used,  $\text{Av. } r = \frac{M \text{ Av. } S^2 - 1}{m - 1}$  (Woodworth: Combining the Results of Several Tests).

The first column stands for the Pearson coefficient obtained

from the formula  $r = \frac{\sum (x,y)}{N \sigma_A \sigma_B}$  or, as it is usually stated,

$r = \frac{\sum (x,y)}{\sqrt{\sum x^2} \sqrt{\sum y^2}}$ . The P. E. in this case means the probable divergence of the true coefficient of correlation from that obtained from a limited random selection of cases. The formula was  $\sigma T\text{-obt. } r = \frac{1-r^2}{\sqrt{n}}$ . If the median deviation of the

probable divergence is desired it may be obtained by multiplying the figures in the second column by .6754. For a discussion of these formulae and any other statistical methods here used, see E. L. Thorndike, *Mental and Social Measurements*.

It is interesting to find the Alpha test correlating most closely with the Stanford-Binet of all the tests used. It corroborates to some extent the current opinion that the Stanford-Binet is largely a test of language ability.

The next highest correlation, that of the Myers Mental Measure, is more difficult to explain. Although intended as a non-language group intelligence test, it involved more language than any of the other tests employed. Still it would seem surprising if this were such a tremendous factor. It tends to indicate the validity of group tests as does also Alpha, in that these two tests were given to nearly all of the subjects in groups, and yet correlate more highly with the individual Stanford-Binet than any of the other tests do, practically all of which were given individually.

One of the most surprising results is the high correlation of tapping with the Stanford-Binet. One would generally assume that the type of motor ability required in our tapping test had little to do with intelligence—especially with older subjects. Our data apparently contradict this hypothesis, and we are confronted with the necessity of explaining the data. It is known that tapping ability increases with chronological age at least up to maturity in the absence of tremors, epilepsy, chorea and other diseases affecting the co-ordinating mechanisms. When we consider that our subjects were all normal and therefore their mental ages tended to increase with their chronological ages, and that all our cases were treated together regardless of age, it at once seems plausible that we have here a spurious correlation due to increase in both scores with chronological age, rather than intelligence.



We have therefore correlated tapping with the Stanford-Binet I. Q.'s, which represent intelligence regardless of age, the coefficient obtained being .069, and find that our assumption is justified.

The correlations between the construction tests and Stanford-Binet are very low, the only reasonably high coefficient being obtained with Healy A. This correlation was about the same as the composite of construction tests with Stanford-Binet. It is interesting to note that Healy A was the only construction test which Professor Terman used in his revision inasmuch as he considered that one only to meet the requirements sufficiently to be included.

The remaining tests, Pintner, Porteus, Learning, and P. C. 2, that have been correlated with Stanford-Binet, each show a correlation very close to .50. This we consider significant in that they are sufficiently high to show that we are measuring intelligence, restricting that term to its generally used meaning with reference to mental testing. In addition the coefficients of correlation are low enough so that we may conclude that different abilities of the subjects tested are being measured, that is, the use of different tests does not result in a repeated measurement of the same abilities. Consequently the use of these tests in addition to the Stanford-Binet, means the measurement of more varieties of ability than can be tested by the Stanford-Binet alone. It remains to be determined whether Pintner, Porteus, Learning, and P. C. 2 all measure the same factor or whether some if not all of them can be used to distinguish special abilities which the others do not test. The answer to this inquiry lies in the results obtainable by the correlation of all of these tests with each other. These results are recorded in Table 16 and they are results so unexpected that they call for interpretation.

Many of the correlations recorded in the table show that the importance of the language factor has been overestimated in dealing with older school children. The correlation between language and non-language tests are high enough to show that the language factor need not be avoided to have a test which can be said to measure intelligence.

Let us first consider the correlation between Myers and Alpha 2; the former is supposed to be a non-language test, the latter a test of understanding of sentences. If language were an important factor it would be hard to account for a

correlation of .733, the second highest obtained aside from the self-correlations. P. C. II is a performance test dealing with pictures; it is concrete where the Alpha 2 deals with abstract ideas,—yet these two give a correlation of .709—likewise unquestionably high. We have stated elsewhere that some language enters into the Myers Mental Measure, and that ability to respond to the spoken word *may* be an exceedingly important factor. If so, why does P. C. II give a coefficient of .714 with Myers Mental Measure? If the Myers Mental Measure is a non-language test, why is the correlation of Pintner, a thorough going non-language test, with Myers lower than Pintner with Alpha 2, a test involving so much language? Again, when we compare the correlation of Alpha 2 and Porteus—.701—with that of P. C. II and Porteus—.702—we are at a loss to explain the similarity in result unless we discard the idea of the importance of the language factor. For the Porteus test requires no language.

One must not overlook the importance of language as a handicap in giving tests to foreigners, etc., but where older school children are being tested it cannot be vital. For in order to succeed in the higher grammar school grades, it is essential that they have a fairly good working knowledge of the English language, and this is all that is needed to succeed with the so-called language tests.

Coming to the selection of a schedule of tests we conclude that:

1. Reading scale Alpha 2 should be included. In the first place because of its high correlations with other tests, the highest of any test with all the others, and also because of special considerations. It must be remembered that Alpha 2 is entirely a reading and writing test and therefore one would not expect so uniformly high a correlation as exists between it and the other tests which are supposed each to be specially adapted toward bringing out certain abilities. The high correlations remind us of Binet's constant contention that intelligence, broadly speaking, can be tested by language tests. This conclusion, however, does not imply that a non-language test cannot serve a like purpose. Our intercorrelations show that there was no reason to avoid language tests inasmuch as they correlated highly with the non-language tests. This is interesting in that the tendency in devising tests is towards making them language tests. For causes of this tendency



we can ascribe first, the simplicity and lack of apparatus inherent in them and second, that the difficulty or ease of the test is far more readily regulated than in the non-language tests. In scrutinizing a test to forecast the results of its use our results seem to show that it is not necessary to dwell upon whether or not the tests involve the use of language.

2. The list of tests selected includes both Myers and Pintner. Myers correlates more highly with every other test than does Pintner, with the exception of their respective correlations with Porteus. As has been stated, both of these tests are valuable and in addition they have the merit of yielding different results.

3. The intercorrelations of P. C. II with such of the other tests as we found to be reliable were sufficiently high to make us believe that this test should be included in our schedule in spite of the uncertainty as to whether it is reliable.

4. Definite judgment upon the learning tests should be reserved for the present. Their highest correlation is under 50 and we have no evidence that they are reliable. Before the learning tests have a right to be so called they must be shown really to measure learning ability; they must also be tested for reliability. It is a question whether learning ability of a given individual is uniform in all fields. The uniformity of learning ability cannot be assumed, for a mere assumption as to the uniformity of motor ability proved to be wrong. (See Perrin, *An Experimental Study of Motor Ability*). If learning ability is found not to be so, combining the various tests may operate to conceal what is valuable in them.

5. The tapping test should be included, in the discretion of the examiner, not because the results can be relied upon to indicate intelligence but because giving this test, which takes only a minute, may disclose latent defects in motor control.

In an attempt to reach some definite conclusion about the construction tests we have made many correlations of different combinations. Healy A is the only test which gives a correlation as high as .40 and that with the Stanford-Binet. Nor does combining the tests raise the correlations, for the four construction tests correlated with Stanford-Binet give a result practically no higher than Healy A alone. It seemed useless to correlate the construction tests with the other tests when they gave such unsatisfactory results with each other, and with the Stanford-Binet. We have no evidence that these four



tests are valuable either as intelligence tests, or for any other purpose.

The Porteus test is one of the most interesting. Since no language enters into the test, one would expect it to correlate more highly with Pintner and Myers than with the Stanford-Binet and Alpha 2. Just the opposite occurs; of the four, by far the highest correlation is with Alpha 2. The correlation of Porteus and P. C. II is practically the same. These three tests all seem to call for one kind of ability. Is it good judgment, common sense ability, planfulness, deliberation, carefulness, foresight, good apperceptions? Probably it contains these and other similar traits. It is the difference between these tests which brings the correlations down to .70, and which causes them to correlate differently with the other tests. Alpha 2 and Stanford-Binet, both requiring language, correlate more highly than Porteus and Stanford-Binet, or P. C. II and Stanford-Binet. Some other factor causes Alpha 2 and P. C. II to correlate considerably higher with Myers than Porteus does. There are always many traits measured by every test, no matter how simple, and the emphasis on the different factors is not always the same for the same test. It varies with the group being measured; their age, sex, education, social selection, etc. Why does Porteus get a higher correlation between his test and Binet's than we do? Partly, at least, because he tested children of all ages, but especially younger ones, whereas ours group themselves closely about a mode, and are older.

We have quite a number of cases which are not completely measured by either the Stanford-Binet or the Porteus tests; that is, they could probably succeed with some harder tests if they were given the opportunity, and this lowers our correlations. The fact that many of Porteus' cases were placed higher rather than lower on his tests than on the Stanford-Binet seems to show that they tended to be poorer in language ability than in planfulness, apperceptions,—whatever one wishes to call it. Our cases on the other hand seem to find no difficulty with the language factor. It is by comparing the results of the same group in different tests, and of different groups on the same test, that most can be learned of what the tests actually do measure. In this study we have the same group measured by many different tests. We find our inter-correlations high in many cases, but nowhere so high that we

feel that the tests are identical. However certain similarities such as the one just discussed between the Porteus, P. C. II, and Alpha 2, were brought to light by this method. Differences such as the striking one between Pintner and Myers have also been observed. If different groups had been used, one would be unable to draw any conclusions regarding the tests for the groups themselves might be responsible for so many of the factors. Again different factors of the tests are brought out by different groups as for instance a younger and older set of children tested with the Pintner non-language survey test would give an entirely different kind of intercorrelation between the separate tests. By this method we can take account of more factors and so interpret our findings with greater accuracy.

There is no evidence that the P. C. test measures apperceptions, that the learning tests measure learning ability, that the construction tests measure ability to use concrete material. On this account, and also because each involves too many incidental, disturbing factors, none of these tests can be considered adequate measures of special abilities or disabilities. Such tests are much needed, and should be constructed so as to measure fundamental, underlying differences in ability. They must be correlated with everything of any possible importance in order to ascertain the degree to which one ability is related to all others. In studying memory we want to know how important a part it plays in reasoning, in mechanical work, etc. We must learn the significance of a good memory for every school study, and for various occupations. If different kinds of memory play important parts in different studies and vocations, this too we must find out. It is a big task, perhaps impossible to carry out at present, but without such information we are tremendously handicapped. The taboo of "faculty" psychology has contributed to lessen activity along these lines, for if you investigate memory you are getting perilously near something obsolete. But very few would deny that there is such a thing as remembering, and all study of memory and its ramifications has yielded interesting and important results.

It has been more or less tacitly assumed in the past that differences in performance are due to differences in the material used rather than to underlying "faculty" differences. This was based upon findings such as those obtained when

TABLE XVI  
INTERCORRELATIONS

	ST-B	Pintner	Myers	Alpha 2	P.C.II	Porteus	Learning	Tapping
ST-B	.439		.686	.757	.541	.536	.491	.604
Pintner		.439	.584	.597	.423	.356	.268	.282
Myers	.439	.584	.686	.757	.541	.536	.491	.604
Alpha 2	.757	.597	.584	.733	.423	.356	.321	.522
P.C. II	.541	.423	.584	.733	.709	.330	.420	.332
Porteus	.536	.356	.733	.709	.709	.702	.290	.382
Learning	.491	.268	.321	.420	.290	.95	.457	.483
Tapping	.604	.282	.522	.332	.382	.483	.252	.805

ST-B. I. Q. with Tapping  $r = .069$ .ST-B. M. A. with Total number of remaining Tests  $r = .5976$ .ST-B. with average Pintner and Myers  $r = .588$ .Healy A with Knox A—Time— $r = .135$ .Healy A with Knox B—Time— $r = -.055$ .Healy A with Knox B—moves— $r = .126$ .Knox A with Knox B—Time— $r = .27$ .Knox A with Knox B—moves— $r = .103$ .Healy A with Healy B—Time— $r = .21$ .Healy A with Healy B—moves— $r = -.001$ .Average Healy A and B with average Knox A and B—Time— $r = .161$ .Average Healy A and B with average Knox A and B—moves— $r = .076$ .

This Table summarizes the results of the correlations of each Test with every other. It should be noted that a correlation was obtained between Stanford-Binet and all other Tests combined the coefficient of .5976 being fairly high.



memory was tested. It was found that a good memory for logical material did not follow from a good memory for nonsense; that being able to remember visually presented facts did not necessarily indicate ability to remember what was heard. The result of these and similar observations has been the development of tests dealing with specific types of material, or—giving up the specific side entirely—tests of general intelligence. Our data seem to indicate that real, underlying differences do exist, if we only know how to get at them. In order to prove this, it is necessary to have a test with omnibus material, all of which is designed to measure a certain type of thing. We shall now proceed to do this.

## A COMBINATION TEST FOR PLANFULNESS

The correlations in table XVI, particularly those obtained between Porteus, Alpha 2, and P. C. II, seem to indicate the possibility of a factor, common to all and largely determining the score on each, which has nothing to do with the material employed, that is, whether a language or non-language test, or the like. We have suggested above several names for this factor,—good judgment, common sense, deliberation, carefulness, foresight, good apperceptions, planfulness, persistence, prudence and mental alertness in meeting a new situation, ability to see the whole of a situation instead of reacting to the most obvious part of it. An attempt was made to investigate it more thoroughly by combining the elements of each test which seemed most specifically to measure it. The selection was made from the Porteus, Myers, Alpha 2, P. C. II, and Stanford-Binet tests. All the tests selected would require about twenty-five minutes to perform, this being a liberal estimate based upon the time limit for each test. Alpha 2 has no definite time limit, but from the writer's experience, ten minutes would seem ample to allow for the parts of the test included in this selection. When all the individual tests had been chosen, they were divided into two sections, and a self-correlation of .763 was obtained with 80 cases. The tests in each group were:

- I. Porteus—year 11 (scored 0, 1, 2) year 12 (scored 0, 1, 2, 3, 4).  
Myers—pages 4. Numbers 3 and 7 (scored each 0, 1).  
P. C. II—pictures 2 and 6 (scored 1 each if OK; otherwise 0).  
Alpha 2, Part II—difficulty 8—number 4 (scored 0, 1).  
Pintner—test 5, numbers 5 and 7 (scored each 0, 1).  
Pintner—test 6, picture 2 pieces 2 and 1 (scored each 0, 1).  
Pintner—test 6, picture 3, pieces 4 and 1 (scored each 0, 1).

---

- II. Porteus—year 10 (scored 0, 1, 2) year 14 (scored 0, 1, 2, 3, 4).

- Myers—page 4. Numbers 5 and 10 (scored each 0, 1).  
P. C. II—pictures 7 and 8 (scored 1 each if OK; otherwise 0).  
Alpha 2, Part II—difficulty 8—number 1 (scored 0, 1).  
Pintner—test 5, number 6 (scored 0, 1).  
Pintner—test 6, picture 2, pieces 4 and 3 (scored each 0, 1).  
Pintner—test 6, picture 3, pieces 2 and 3 (scored each 0, 1).  
Stanford-Binet—XIV years, number 6 (scored 0, 1).
- 

The Porteus tests were chosen because they were devised to measure this very thing. The fact that only one type of material—mazes—was included, was considered by Porteus one of the outstanding advantages of his test. We feel that this is a disadvantage since some children might have a disability for working with this kind of material although possessed of common sense, foresight, etc. With omnibus material this special factor is overcome. The choice of the four most difficult tests was largely a matter of the distribution of the subjects. Too many would have made perfect records on the easier tests.

The selection from Myers Mental Measure was based largely upon resistance to suggestion. In each case four pictures with some element in common must be chosen from eight possible ones and underlined. These four could not be too difficult or our subjects would all score 0; if they were too easy we would have no reason to believe that this characteristic pertained to them. Number 3 is the selection of four toys,—a tricycle, top, kite and rocking horse, with a soldier as the confusing picture. In number 5, four items made of iron must be chosen,—a stove, dagger, or sword, train, and lock. This has several confusing suggestions. There is a broom which might be associated with the stove, and two animals which might be connected with the train as they all are capable of locomotion. Number 7 consists of an insect, a broom, a bird, a table, a butterfly, an aeroplane, a goat and a cow. The four things which can travel in air are to be underlined. The two animals prove confusing to many children. In number 10 the subject is to select four articles of wood,—two trees, a barrel and a table, with a snake, a camel, a cannon, and a



bird to be omitted. Here also the three animals receive considerable attention, the hasty child not noticing that the fourth is lacking, or the snake is overlooked, the two remaining animals and two trees being classed together as objects possessing life. There seems to be some suggestion in each of these pictures, and it is certainly true that a careful, deliberate, performance by a subject who takes in the whole situation and responds to it will give far better results than a hasty, careless one.

The pictures from P. C. II are those in which there are several obvious possibilities. A hasty, careless selection will hit upon the first possible one, rather than searching further for the exactly correct one. All correct pieces were checked by asking the subject why that particular one had been chosen and if it was put in by chance, no credit was given. The partial credits given by Healy were omitted, the picture scored either as perfect or a failure. This was necessary in order to eliminate the other possible factors which enter into solving the test partly. For instance, in the second picture, where a book is missing, it is not sufficient to put in any book, pencil case or lunch box, but by following up persistently all the clues, the one and only correct red book can be placed in the space with certainty.

From the Thorndike Alpha 2 reading scale questions were selected which had been answered by a large number of children. Question I requires a fairly careful study of the paragraph in order to find just what it is that seems true at first but is really false. The question is a little clumsily put,—certainly not direct and to the point,—which is an advantage for our purposes. Question 4 is not a reading scale problem proper, but necessitates close attention to several directions. In two rows of digits the subject must underline every five that comes just after a two, unless the two comes just after a nine. If that is the case, he must draw a line under the next figure after the five. The last few lines of the first page of the Myers Mental Measure are similar to this, but the Alpha 2 was given to a larger number of cases, there was no time limit, and less possibility of copying, so it was given the preference, as being more accurate.

Numbers 5, 6, and 7 from Pintner test 5 are all similar in nature. Given a drawing, the problem is to draw it in a reversed position, with two lines of the second position given

on which to construct the rest. This seems like a rather special ability, but Pintner gives each drawing considerable weight in his total score, and persistence and planfulness are certainly essential for a good performance.

Pintner test 6 consists of parts of pictures presented in a disarranged order. Each part is numbered and blank spaces are provided in which the subject is to place the numbers of the parts in order which would give a perfect ensemble. Here again planfulness, patience, and foresight are needed, and on the whole the subject who possesses them to the greatest degree will be the most successful.

Finally one test was selected from the Stanford-Binet scale, —namely the reversed clock hands of year XIV. If two out of three were correct a score of one was given, if less no credit at all. This test seemed to require the same kind of ability as many of the other tests included, and was therefore added. Some of the other Stanford-Binet series might have been used also, but those which seemed desirable came too high or too low in the scale so that the distribution for our subjects would not be satisfactory.

The correlation of .763 obtained between the two parts is fairly high when it is remembered that the highest score on each section can only be 17; also that the whole series of both parts would only take half an hour to give. As to reliability it is a noteworthy conclusion that this self-correlation is the highest one obtained with any non-identical material. A correlation of the composite tests with any of the tests which are included would probably give a high coefficient difficult to interpret because of the varying amount of each included in the composites, and a low correlation with learning tests, construction tests, or tapping could hardly be considered strong evidence in favor of our new grouping. But the correlation with Stanford-Binet seemed worth finding, and when worked out yielded a coefficient of .537. This indicates that our combination test is comparable with the whole series of tests from which it was compiled. We have, however, no criterion to prove that it actually measures the trait which we presuppose it does. But this same criticism applies to all the tests which are supposed to measure specific factors. Our new test combination of old material is certainly as good as the tests from which it originated; we think it is better, because it gives evidence of measuring one trait, or group of

traits with a variety of materials, whereas all the others measure many kinds of traits with identical or similar material. That is, the classification and material preparatory to the formation of a test has generally heretofore been along the lines of the material employed, such as form boards, etc., whereas the combination test being discussed presents the results obtained from forming a test directed toward planfulness, or other ability.



## CONCLUSION

It is proposed to set forth the practical results of this study, to show the positive information that has been ascertained and also to show from the experience gathered in the course of obtaining such information, what further investigations should be made, with what purpose, and what methods may lead to success. This study has reached some positive results and has disclosed other perhaps more valuable ones in the same field.

In entering upon this study it was believed that the results of the method that has been pursued would justify the conclusion that the Stanford-Binet series can be used as a test of general intelligence and that certain other tests used as auxiliaries would make apparent and give a measure of special abilities not individually measured by the Stanford-Binet. It was expected that the various tests would give reasonably high correlations with the Stanford-Binet and rather low correlations with each other, thus on the one hand establishing the reliability of the tests used, and on the other hand, the diversity of the abilities that were subjected to measurement.

These results were anticipated because care was used in selecting the tests to take those which had an approved authorship, an extended use, a definite purpose, and a general reputation of success in the field they purported to cover. That is, the various units had each been shown apparently to be satisfactory and on these *a priori* grounds it was thought that properly selected units used in conjunction would result in a reliable schedule.

Had the results of the correlations been in harmony with this anticipated situation, we might properly have pointed to this study as a demonstration of the process by which schedules of tests for children should be composed.

Looking upon our results as they have been reported upon, the fact is obvious that there is no such easy manner in which to arrive at reliable schedules of tests. Unexpected low correlations were obtained in some situations where the indicated results should have been high, and vice versa, and while our positive purpose therefore met with disappointing obstacles, a study of the figures as we have them led to other worthwhile conclusions.

Drawing upon the results of the correlations, it can be stated with assurance that it will not be well to take tests upon which a high face value has been placed when they were used without being effectively valued by comparison, and combining a number of them in the expectation of using the combination to get reliable information as to the general intelligence and the special abilities of normal children. One of the best examples which we can show, as a result of this study, of the impropriety of such procedure is, that the type of material used does not govern the abilities tested. We obtained a higher correlation between a language and a non-language test than between two language tests or two non-language tests, similar examples can be drawn from the correlations listed above respecting other characteristics of various tests. Insofar therefore as authors of tests have relied upon the material as a quality that would single out and measure a certain one of many abilities, it seems clear that individual tests miss their purpose. However, the correlations did seem to show that something definite was being tested, so that if our purpose of finding a schedule of tests at once sufficient to measure both general and special abilities, was disappointed, at least the schedule we used can be relied upon for general abilities and that such a schedule is more reliable than the Stanford-Binet alone. The components of this schedule have been previously listed and it only remains to state what individual matters of interest relating to each were made clear in the course of the study which was directed to larger purposes.

It was a matter of actual demonstration herein that all of the construction tests used are unreliable, this conclusion disproving the previously held opinion based upon empirical considerations to the effect that they reliably measure ability to handle concrete material.

Persons having occasion to apply mental tests have too frequently overlooked the matter of how far the test can be relied upon. This is an important matter and consequently it should be of some interest to note that the reliability of the Stanford-Binet, Pintner non-language group test, Thorndike reading scale Alpha 2, Porteus Maze Test, and tapping test has been established, whereas the Myers Mental Measure, the Healy Pictorial Completion test II, the Healy-Bronner

learning tests and the crossline tests are not yet definitely shown to be reliable.

Care should also be observed in interpreting the results of correlations, for the mere fact of high correlation is only generally and not conclusively proof of reliability. There is the possibility that factors causing unreliability have been hidden—thus, in the tapping test, the high correlation with Stanford-Binet was deceptive owing to the fact that the scores on both increased with the age of the subjects. Other specific remarks relating to individual tests are contained in the results.

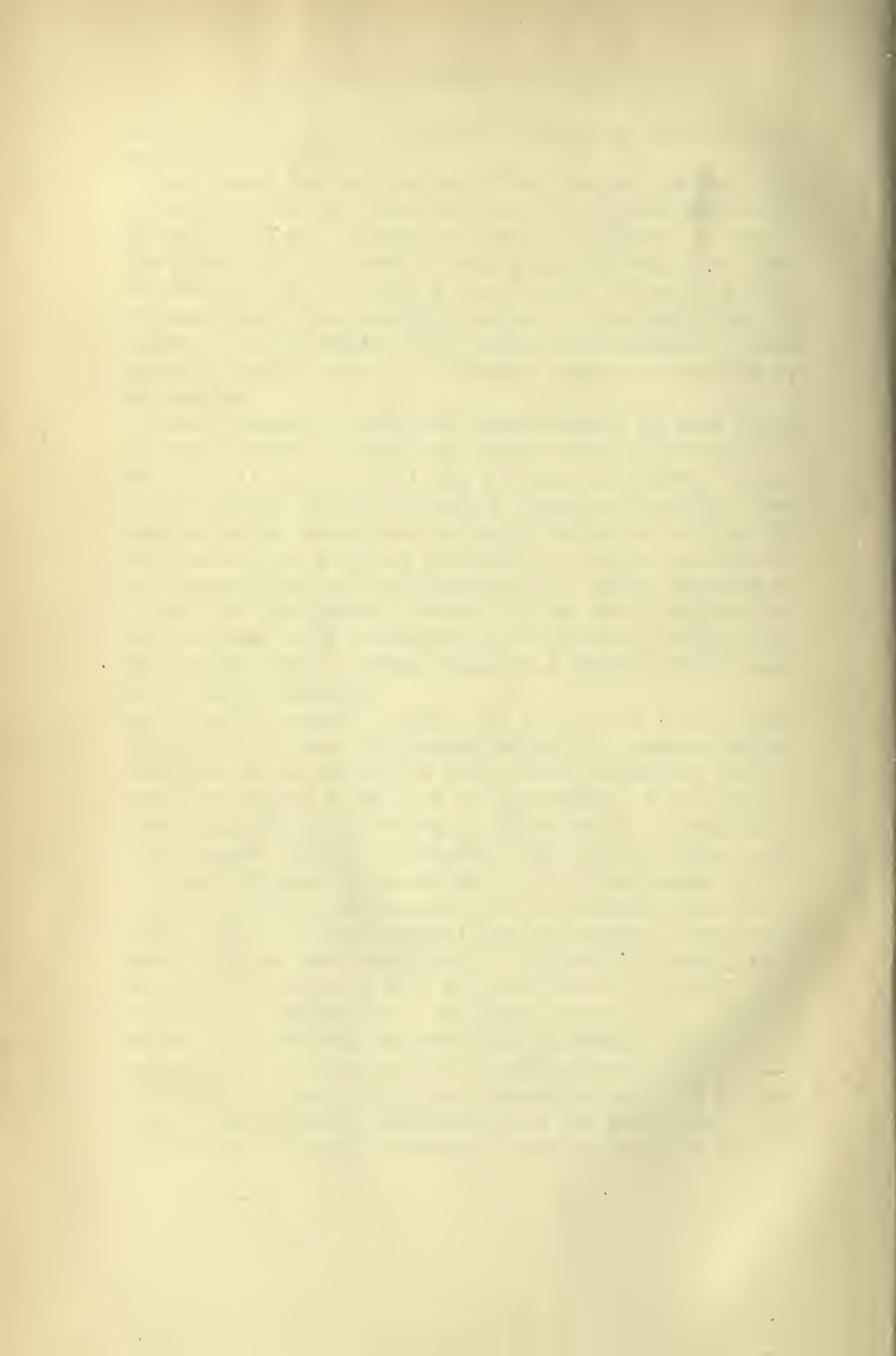
There remains to state what considerations we have found to have a probable value as to future work in this field. If we found on the one hand that the type of material used in a test does not govern the ability tested, on the other hand there are some indications that to test individual abilities the test should have a variety of material. So far the elements of a desired test can be stated, but the further necessity of finding just what material is suitable, can only be determined by practical work consisting of correlation with outside criteria and with any other measures of claimed effectiveness in the field in question.

As an experimental example, for the confines of this study would allow no more extended investigation, various parts of a number of the tests were united in a combination test intended to secure a measure of planfulness. The resulting correlations indicated success in this attempt. A similar or even greater measure of success may follow further combinations aimed at the measurement of other abilities.

It may also be stated as having been illustrated in the course of this study that the supposed merit of various mental tests based upon various, insufficient or unscientific criteria, such as mere hypothesis, or even practical results, if relied upon, may lead to misleading or dangerous conclusions, and that before one takes the responsibility of giving advice or of taking action with respect to information gained from the application of mental tests, there should be available the assurance that proper comparative tests and correlations have verified the supposed propriety of relying upon the results.

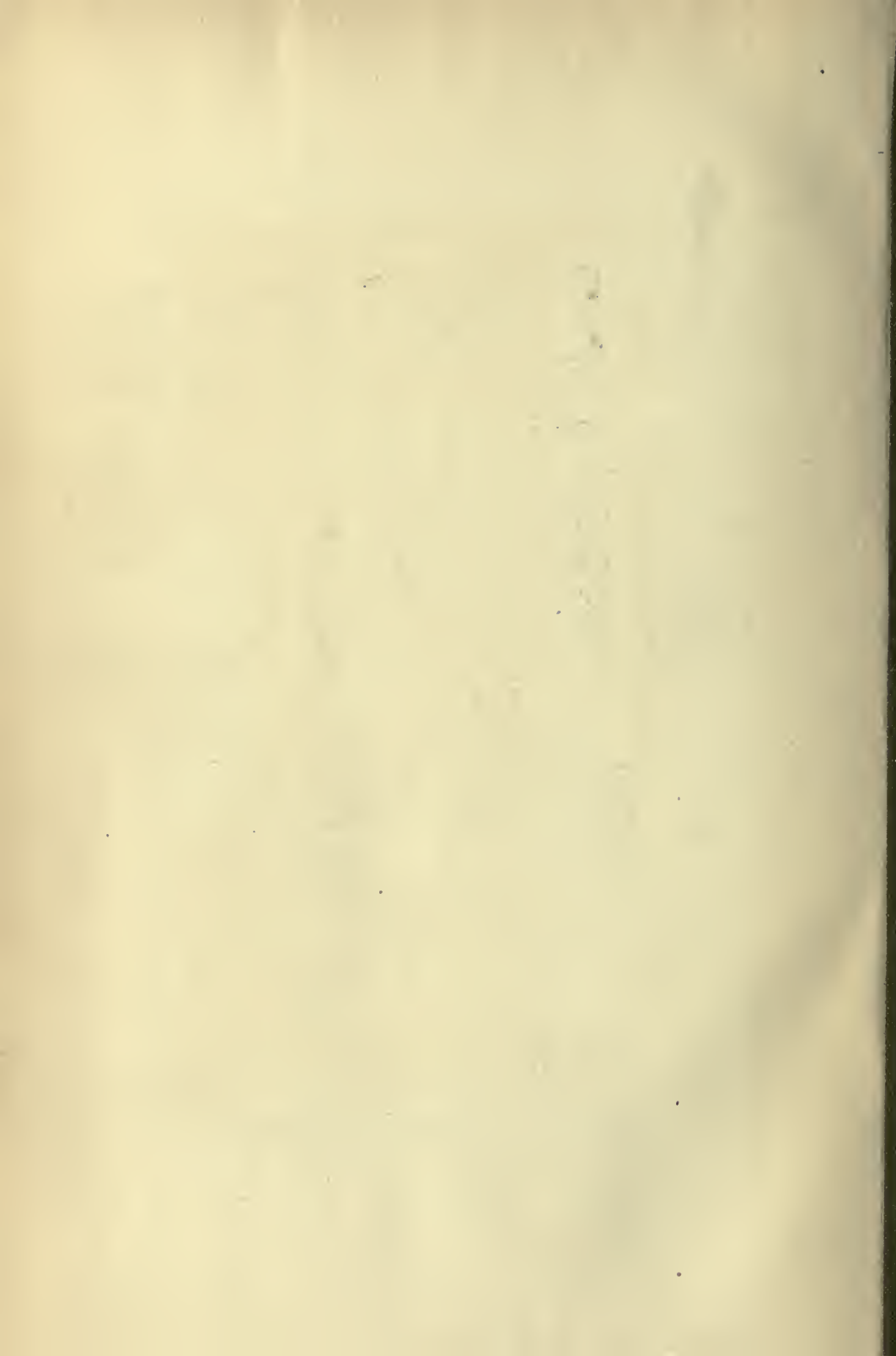












BINDING LIST DEC 1 1929

CIRCULATE AS MONOGRAPH

BF  
21  
A7  
v.7

Archives of psychology

PLEASE DO NOT REMOVE  
SLIPS FROM THIS POCKET

---

---

UNIVERSITY OF TORONTO  
LIBRARY

